



(12) 发明专利申请

(10) 申请公布号 CN 103124784 A

(43) 申请公布日 2013. 05. 29

(21) 申请号 201180041130. 4

(22) 申请日 2011. 08. 23

(30) 优先权数据

61/376, 122 2010. 08. 23 US

(85) PCT申请进入国家阶段日

2013. 02. 25

(86) PCT申请的申请数据

PCT/US2011/048819 2011. 08. 23

(87) PCT申请的公布数据

W02012/027376 EN 2012. 03. 01

(71) 申请人 纽约州立大学研究基金会

地址 美国纽约

(72) 发明人 尤珀·玛

(74) 专利代理机构 上海天翔知识产权代理有限公司

公司 31224

代理人 刘粉宝

(51) Int. Cl.

C12N 5/074 (2006. 01)

C12N 5/0789 (2006. 01)

C12N 5/02 (2006. 01)

C07K 14/47 (2006. 01)

权利要求书2页 说明书31页

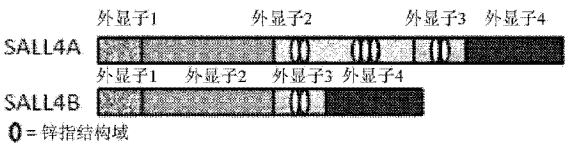
序列表34页 附图37页

(54) 发明名称

用于扩增干细胞的方法和这些细胞的用途

(57) 摘要

本发明证实, SALL4A 和 SALL4B 是造血干细胞扩增的较强正调控因子。接受 SALL4A 或 SALL4B 的表达的 HSC 能够实现高水平扩增。用 SALL4 转导的细胞的培养物引起广泛的 HSC 扩增, 其扩增水平在 2 到 3 周内达到对照物的超过 1000 倍, 并且扩增的 HSC 未显示或显示极少成熟。而且, 所述扩增的发生相当迅速, 仅在数天内就出现显著 HSC 生长。此外, SALL4 诱导的 HSC 扩增未削弱造血细胞的分化。对于 HSC 来说, SALL4 看起来起到维持未分化的增殖状态的作用并且阻止细胞分化。



1. 一种用于扩增干细胞群的方法,所述方法包含向所述干细胞群提供有效扩增所述干细胞群的量的多肽,所述多肽具有 Sal 样(SALL)多肽的增强扩增活性。

2. 根据权利要求1所述的方法,其中所述干细胞是成体干细胞。

3. 根据权利要求1或2所述的方法,其中所述干细胞是在以下各项中或来源于以下各项:脑、肝、心脏、肾、皮肤、胰腺、膀胱、胆囊、大肠、小肠、胃、骨骼肌或肺。

4. 根据权利要求1或2所述的方法,其中所述干细胞是造血干细胞。

5. 根据权利要求4所述的方法,其中所述造血干细胞是在以下各项中或来源于以下各项:脐带血、外周血、骨髓或脾。

6. 根据权利要求1到4中任一权利要求所述的方法,其中所述造血干细胞是人类干细胞。

7. 根据权利要求1到6中任一权利要求所述的方法,其中所述 SALL 多肽连接到转运部分,所述转运部分能够跨过细胞膜,由此将所述 SALL 多肽转运到所述细胞中。

8. 根据权利要求7所述的方法,其中所述转运部分是 HIV-1 转录转活化因子(TAT)肽、Chariot 蛋白、富精氨酸肽、触足源性穿膜肽、1 型单纯疱疹病毒 VP22 蛋白或 +36GFP。

9. 根据权利要求1到8中任一权利要求所述的方法,其中所述 SALL 多肽包含如 SEQ ID No:7、8、9、10、11 或 12 所陈述的氨基酸序列。

10. 根据权利要求1到9中任一权利要求所述的方法,其中所述干细胞群扩增 10 倍、20 倍、50 倍、100 倍或 1000 倍。

11. 根据权利要求1到10中任一权利要求所述的方法,其中所述干细胞群是离体扩增。

12. 根据权利要求1到11中任一权利要求所述的方法,其中所述细胞群是在包含 50ng/ml 的 FMS 样酪氨酸激酶-3 (FLT-3)、50ng/ml 血小板生成素(TPO)和 / 或 50ng/ml 干细胞因子(SCF)的培养基中培养。

13. 根据权利要求11或12所述的方法,其中所述干细胞群是在包含 25ng/ml 的 FMS 样酪氨酸激酶-3 (FLT-3)、25ng/ml 血小板生成素(TPO)和 / 或 25ng/ml 干细胞因子(SCF)的培养基中培养。

14. 根据权利要求11到13中任一权利要求所述的方法,其中所述 SALL 多肽是由包含如 SEQ ID No:1、2、3、4、5 或 6 所陈述的序列的核苷酸序列所编码。

15. 根据权利要求11到14中任一权利要求所述的方法,其中所述干细胞群具有包含如 SEQ ID No:7、8、9、10、11 和 / 或 12 所陈述的氨基酸序列的 SALL 多肽。

16. 根据权利要求11到15中任一权利要求所述的方法,其中所述群体中的干细胞用包含编码所述 SALL 多肽的核苷酸的病毒载体转导,由此将所述 SALL 多肽提供给所述干细胞群。

17. 根据权利要求16所述的方法,其中编码所述 SALL 多肽的核苷酸包含如 SEQ ID No:1、2、3、4、5 或 6 所陈述的核苷酸序列。

18. 根据权利要求16所述的方法,其中编码所述 SALL 多肽的所述核苷酸是在诱导型启动子的控制下表达。

19. 一种用于增进受检者体内干细胞群的扩增的组合物,所述组合物包含有效扩增所述干细胞群的量的 SALL 多肽,和培养基。

20. 根据权利要求19所述的组合物,另外包含干细胞。

21. 一种鉴别用于扩增干细胞群的试剂的方法,所述方法包含 (a) 获得候选试剂 ;(b) 将来自所述群体的干细胞暴露于所述候选试剂 ;和 (c) 确定在所述干细胞中 SALL 多肽是否上调,其中如果 SALL 多肽上调,那么确定所述试剂是用于扩增所述干细胞群的试剂。

22. 一种 SALL 多肽的用途,其用于离体干细胞群,包含向所述干细胞群提供有效离体扩增所述干细胞群的量的所述 SALL 多肽。

23. 一种 SALL 多肽的用途,其用于治疗或预防需要干细胞或由所述干细胞得到的扩增的干细胞的受检者的疾病、病症或异常,包含 a) 获得干细胞群 ;b) 向所述干细胞群提供有效扩增所述干细胞群的量的所述 SALL 多肽 ;和 c) 将有效治疗或预防所述受检者的所述疾病、病症或异常的量的所述扩增的干细胞群移植到所述受检者。

24. 一种制造用于扩增干细胞群的试剂的方法,包含 :

a) 获得候选试剂 ;

b) 将来自所述群体的干细胞暴露于所述候选试剂 ;

c) 确定在所述干细胞中 SALL 多肽是否上调,其中如果 SALL 多肽上调,那么确定所述试剂是用于扩增所述干细胞群的试剂 ;及

d) 选择步骤 c) 中已经确定为用于扩增所述干细胞群的试剂的试剂。

25. 一种用于治疗或预防需要干细胞或由其得到的扩增的干细胞的受检者的疾病、病症或异常的方法,所述方法包含 a) 获得干细胞群 ;b) 向所述干细胞群提供有效扩增所述干细胞群的量的 SALL 多肽 ;和 c) 将有效治疗或预防所述受检者的所述疾病、病症或异常的量的所述扩增的干细胞群移植到所述受检者。

26. 一种干细胞库,包含遗传上不同的干细胞群,其中所述干细胞群已经根据权利要求 1 所述的方法扩增。

27. 一种用于扩增受检者体内的干细胞群的方法,包含增加所述受检者体内的 SALL 多肽量。

28. 一种用于增加受检者体内祖细胞群的规模的方法,所述方法包含增加所述受检者体内的 SALL 多肽量。

29. 根据权利要求 27 或 28 所述的方法,其中增加所述受检者体内的 SALL 多肽量包含将所述 SALL 多肽引入所述受检者体内。

30. 根据权利要求 27 或 28 所述的方法,其中增加所述受检者体内的 SALL 多肽量包含将核酸分子引入所述受检者体内,所述核酸分子具有编码所述 SALL 多肽的核苷酸序列。

31. 根据权利要求 30 所述的方法,其中编码所述 SALL 多肽的所述核苷酸包含如 SEQ ID No:1、2、3、4、5 或 6 所陈述的核苷酸序列。

32. 根据权利要求 27 到 31 中任一权利要求所述的方法,其中所述 SALL 多肽包含转运部分,所述转运部分能够跨过细胞膜,由此将所述 SALL 多肽转运到所述细胞中。

用于扩增干细胞的方法和这些细胞的用途

[0001] 本申请案主张 2010 年 8 月 23 日申请的美国临时申请案第 61/376,122 号的优先权,该案内容以引用的方式并入本文。

[0002] 获得本发明的研究部分地得到了美国国立卫生研究院(National Institutes of Health)所授予的基金项目号 HL087948 的支持。因此,美国政府在本发明中享有某些权利。

[0003] 在本申请案中,提到的各个出版物都是以括号中的阿拉伯数字表示。有关这些出版物的完整引用提供于最后的参考文献部分中。参考文献部分中引用的出版物的公开内容以全文引用的方式并入本申请案中,以便更完整地描述到本文中描述的方法和设备之日时的技术现状

技术领域

[0004] 本发明涉及使用具有 Sal 样(Sal-like, SALL)多肽的增强扩增活性的多肽离体扩增干细胞群,以及这些细胞的用途。

背景技术

[0005] 在早期的生命和生长期间,干细胞能够在体内发育成许多不同的细胞类型。干细胞可以分为两大类:胚胎干细胞和成体干细胞。在正在发育的胚胎中,干细胞可以分化成所有特种胚胎组织。在成年生物体中,干细胞和祖细胞充当身体的修复系统,补充特种细胞,而且还维持例如血液、皮肤或肠组织等再生器官的正常周转。成体干细胞可以按多种路径分化。

[0006] 间充质干细胞是产生多种细胞类型的成体干细胞:骨细胞(bone cell/osteocyte)、软骨细胞(cartilage cell/chondrocyte)、脂肪细胞(fat cell/adipocyte)以及其它种类的结缔组织细胞,例如腱中的细胞。神经干细胞是脑中的成体干细胞,它们产生三个主要细胞类型:神经细胞(神经元),以及两类非神经元细胞—星形细胞和少突胶质细胞。上皮干细胞是在消化道内衬中的成体干细胞,这些干细胞出现在隐窝深部中并且产生数种细胞类型:吸收细胞、杯状细胞、潘氏细胞(paneth cell)和肠内分泌细胞。皮肤干细胞是存在于表皮的基底层中以及毛囊基部处的成体干细胞。表皮干细胞是产生角质细胞的成体干细胞,这些角质细胞迁移到皮肤的表面并形成保护层。毛囊干细胞可以产生毛囊和表皮两者。

[0007] 造血干细胞(Hematopoietic stem cell, HSC)是稀少的成体干细胞,这些干细胞已经在胎骨髓、胎肝、脐带血、成体骨髓和外周血中鉴别到,它们能够分化成三个细胞谱系,包括髓红样细胞(红细胞、粒细胞、单核细胞)、巨核细胞(血小板)和淋巴样细胞(T细胞、B细胞和自然杀伤细胞)。这些 HSC 被用于临床移植手术中以治疗多种疾病,包括恶性和非恶性病症。HSC 的扩增具有重要的临床应用,因为在离体扩增造血干细胞方面的相对无能极大地限制了目前 HSC 移植术的使用。用于治疗骨髓移植或遗传性病症患者的 HSC 存在短缺。对于异基因骨髓移植来说,在可能从 HSC 移植获益的所有患者中仅有三分之一的人能找到适合的人白细胞抗原(human leukocyte antigen, HLA)匹配相关的供体。

[0008] 在可用干细胞数量有限的情形中更是如此。这包括供移植到成人体内以及输注多个脐血单位的脐血源性干细胞。尽管这些程序可能会有效增加定植的总发生率,但它们无法克服造血恢复速度缓慢的问题。在脐带血移植(umbilical cord blood transplantation,UCBT)后髓样细胞定植的延迟通常伴随着早期移植相关性发病率和死亡率增加。这一点仍是成功地使用脐血作为干细胞的替代来源用于异基因移植的主要障碍,并且需要新颖的策略来克服这一问题。

[0009] 骨髓干细胞已经被用于治疗多种疾病:白血病、炎症、免疫疾病、血液和免疫系统的先天性异常、再生障碍性贫血和血红蛋白病。然而,找到一个匹配的供体是很难而且相当耗时的。三个患者中仅有一个能找到适合的供体,并且许多患者因无法找到适合的供体而死亡。此外,对于非裔美国人、拉丁裔美国人、土著美国人以及混合民族的人来说,找到一个适合的配型尤其成问题。因此,需要开发出一种用于生长造血干细胞的方法,这种方法将解决对于人类供体的需求。建立一个含有不同骨髓干细胞单倍型的细胞库可以使来自一位供体的细胞能够产生足够的供应量用于超过 1,000 个接受者。

[0010] 干细胞还可能掌握着对抗 HIV 的关键。处理血液细胞以使其对 HIV 感染具有抗性的可能方法包括以基因方式改变可分化成 T 细胞的干细胞上的受体。随后,可以使经过修饰的干细胞扩增,并引入 HIV 患者体内。

[0011] 目前,HSC 的标准来源是骨髓和外周血。为了获得骨髓细胞,供体必须经历多次抽取,以收集数千毫升的骨髓,这一程序是在全身麻醉下进行的。为了从外周血中收集到 HSC,供体必须用粒细胞集落刺激因子处理,以增加循环的 HSC 的数量。这两个程序存在一定风险并产生大量成本。

[0012] HSC 的一个较新的重要来源是脐带血(umbilical cord blood,UCB)。脐带血具有优于例如来自骨髓和动员的外周血等其它 HSC 来源的重要优势。UCB 不仅易于从美国近 50 个 UCB 库中的许多库获得,它还对于与宿主主要组织相容性复合物(major histocompatibility complex,MHC)的错配显示出增加的耐受性。

[0013] 除相对广泛的可用性外,这些 HSC 还具有数个有用特性,包括其诱导免疫反应性的能力降低。在许多情形中,与其它 HSC 来源相比较,UCB 的使用显著减少了移植物抗宿主疾病。

[0014] 然而,尽管使用 UCB 的益处很明显,但仍存在一些关键问题,这些问题构成了扩大使用这一造血干细胞来源的主要障碍。成功地使用脐带血作为干细胞来源用于异基因移植的一个阻碍是髓样细胞定植的延迟。这将导致在输入脐带血后早期移植相关性发病率和死亡率增加。尽管有密集并且昂贵的支持性护理,但在移植后的头 100 天期间,治疗相关性死亡率仍大于 50%,这是因为免疫系统和血小板恢复的延迟,使患者易于发生机会性感染。输注多个脐血单位已经被用作增加总体定植的一种可能方法,但到目前为止,尚未解决造血恢复缓慢的问题。

[0015] 扩大 UCB 的使用的另一障碍是采集时每一脐带的 HSC 数量有限。由于细胞剂量已经被证实为 UCB 移植后定植和存活的主要决定因素,故干细胞数量少是成功进行 UCB 干细胞移植的最主要障碍。

[0016] 在移植前离体扩增单个脐带血单位的干细胞组分的能力将大幅增加这一治疗方式的可行性。与未扩增的脐血单位中可用的细胞有限相对,对患者输注较大量的干细胞将

会大幅增加成功定植的可能性。

[0017] 非造血成体干细胞,包括从例如脑、心脏、肝、胰腺、肾、肺等器官分离的干细胞的扩增具有重要的临床应用,特别是用作外部细胞来源以补充组织或器官失去或受损的细胞。

[0018] 此外,用于血液科遗传性病症的干细胞基因疗法因基因转移到早期造血祖细胞和干细胞中的效率很低而受到限制。需要克服的障碍是扩增经过基因修饰的细胞群,从而可以在施用于人类之前获得足够的经过修饰的细胞。举例来说,可以用骨髓移植来治愈患有严重镰状细胞疾病的儿童。在镰状细胞疾病的情形中,无需完全破坏接受者的骨髓,而仅仅是将其用足够健康或经过遗传校正的干细胞替代,由此产生足够数量的健康红细胞。

[0019] 造血干细胞(HSC)的扩增仍然是开发针对骨髓移植和许多血液病症的先进细胞疗法的重要目标。在过去二十年间,由于鉴别出了首批造血生长因子,故已经多次尝试使用经过纯化的已知会调控HSC的生长因子在体外扩增HSC。然而,这些尝试的成功次数有限。举例来说,造血生长因子胎肝酪氨酸激酶(fetalliver tyrosine kinase, Flt3)配体、干细胞因子以及白细胞介素6和11促进鼠类造血干细胞的自更新。然而,与新鲜的输入细胞相比较,仅观察到有限的造血干细胞扩增(1-3)。

[0020] 尽管鉴别出了许多的多能胚胎干(embryonic stem, ES)细胞基因,但都不能作为HSC扩增的稳固因子。如文献中所报导的,它们在HSC扩增中展现出有限的作用或未展现作用。到目前为止所报导的有关多能基因的最佳研究是OCT4和Nanog,并且这两者都无法诱导HSC扩增。这一结论也得到了本研究的支持:在使用病毒载体对这些基因进行强迫表达的情况下,这些基因在组织培养中对HSC扩增不存在显著作用。

[0021] 已经研究将细胞固有路径中的Notch-1活化作为一种增加HSC扩增的可能方式,并且这些研究已经显示,活化这些路径能够维持HSC的淋巴髓质再群体化(repopulation)潜能。HOXB4的过表达是到目前为止所报导的最有效的干细胞扩增方法。近来,安托查克(Antonchuk)等人显示,HOXB4在体外持续10到14天的逆转录病毒过表达可以使再群体化HSC的数量相比新鲜骨髓干细胞增加40倍(4)。

[0022] 然而,即使再群体化HSC增加40倍也不足以实现多个目的。

发明内容

[0023] 本发明公开一种使用Sa1样(SALL)多肽扩增干细胞群的方法。

[0024] Sa1样(SALL)家族(又称Hsa1),包含SALL1转录变异体1(SEQ ID No:1[NCBI参考序列:NM002968.2]),SALL1转录变异体2(SEQ ID No:2[NCBI参考序列:NM001127892.1]),SALL2(SEQ ID No:3[NCBI参考序列:NM_005407.1]),SALL3(SEQ ID No:4[NCBI参考序列:NM_171999.2]),SALL4a(SEQ ID No:5[GenBank:AY172738.1])和SALL4b(SEQ ID No:6[GenBank:AY170621.1]),最初都是基于与果蝇基因sa1同源的DNA序列克隆的。在一个相关方面中,包含如SEQ ID No:1、2、3、4、5和6所陈述的序列的核酸序列分别编码如SEQ ID No:7、8、9、10、11和12所陈述的氨基酸序列。

[0025] 在人体中,SALL家族的成员(包括SALL4)在正常发育中起到重要作用。除在发育中的重要作用外,还发现SALL基因家族在人类和鼠类ES细胞中以及在早期发育过程中有表达。SALL4是在胚胎的两个细胞期中表达,类似于OCT4,而SOX2和NANOG的表达在胚胎

发育的囊胚期中开始(1-3)。本研究组和其它团体已经显示,胚胎干细胞(ESC)因子 SALL4 在维持 ES 细胞多能性以及通过 Oct4 和 Nanog 的转录调节来控制影响 ES 细胞命运的决定方面起到至关重要的作用(4, 6, 8-10)。我们和其它团体也已经显示, SALL4 可以活化 OCT4 并与 Nanog 相互作用(9-11),并且 SALL4/OCT4/Nanog 转录核心网络是维持 ES 细胞的“干细胞性(stemness)”必不可少的。到交配后 10.5 天时, SALL4 主要在各种器官系统(包括胚胎的脑和骨髓)中的干细胞/祖细胞群中可检测到,并且稍后会在成体中检测到。这可以表明, SALL4 不仅涉及于 ESC 中,而且还涉及于成体干细胞中(12-14)。

[0026] 本发明提供一种用于离体扩增干细胞群的方法,所述方法包含向所述干细胞群提供有效离体扩增所述干细胞群的量的多肽,所述多肽具有 Sa1 样(SALL)多肽的增强扩增活性。

[0027] 本发明另外提供一种用于增进受检者体内干细胞群的扩增的组合物,所述组合物包含有效扩增所述干细胞群的量的 SALL 多肽,和培养基。

[0028] 本发明另外提供一种鉴别用于扩增干细胞群的试剂的方法,所述方法包含(a)获得候选试剂;(b)将来自所述群体的干细胞暴露于所述候选试剂;和(c)确定在所述干细胞中 SALL 多肽是否上调,其中如果 SALL 多肽上调,那么确定所述试剂是用于扩增所述干细胞群的试剂。

[0029] 本发明另外提供一种用于治疗或预防需要干细胞或由其得到的扩增的干细胞的受检者的疾病、病症或异常的方法,所述方法包含 a) 获得干细胞群;b) 向所述干细胞群提供有效扩增所述干细胞群的量的 SALL 多肽;和 c) 将有效治疗或预防所述受检者的所述疾病、病症或异常的量的扩增的干细胞群移植到所述受检者。

[0030] 本发明另外提供一种干细胞库,包含遗传上不同的干细胞群,其中所述干细胞群已经根据权利要求所述的方法扩增。

附图说明

[0031] 图 1. SALL4A 和 SALL4B 同功异型物的示意图,描绘了各自所具有的不同数量的锌指结构域。HSC 是使用慢病毒转染系统,用 SALL4A 或 SALL4B 基因进行转导。

[0032] 图 2. 人骨髓 CD34+ 细胞的明视场图像和荧光图像。明视场图像(左图)和荧光图像(右图)说明了含有 GFP+SALL4A、GFP+SALL4B 以及单独 GFP 的慢病毒构建体的感染效率。图像是在用慢病毒感染后 48 小时取得。

[0033] 图 3. (A) 在慢病毒感染后 7 天,用 SALL4A 和 SALL4B 转导的 HSC 能够存活并迅速扩增。(B) 从来自 3 位不同患者的外周血干细胞分离的 CD34+ 细胞。CD34+ 细胞是使用磁性抗 CD34+ 人类微珠从干细胞池分离。富含 CD34+ 的细胞用 SALL4A 转导,并在明视场和荧光显微镜下成像。来自各个患者的全部三份样品都用 SALL4A 成功地转导并且在培养中迅速扩增。此外,当生长因子浓度降低时, SALL4 诱导的 HSC 能够扩增。

[0034] 图 4. 感染后 9 天用 GFP 转导的人骨髓 CD34+ 细胞(i 和 ii)或代表性 SALL4 同功异型物 B(v 和 vi)的明视场图像和荧光图像。最初,接种 50,000 个 CD34+/CD38- 细胞。高放大倍率的用 SALL4B 转导的 HSC 簇(vii 和 viii)。GFP 细胞簇预示着 SALL4B 的阳性过表达。在 SALL4B 过表达的情况下, HSC 细胞簇在感染后 9 天能够存活并迅速扩增。

[0035] 图 5. SALL4 诱导的 HSC 与对照细胞的细胞扩增的比较。在慢病毒感染后 2 天时,

在 GFP 转导的细胞培养物和 SALL4 转导的细胞培养物中可见类似量的 HSC 克隆。感染后 11 天时, SALL4 转导的 HSC 增殖并扩增。注意到与 GFP 诱导的对照 HSC 中细胞数量较少相比较, 在过表达 SALL-4 的 HSC 中形成正在增殖的细胞簇。

[0036] 图 6. 慢病毒感染后 16 天, SALL4 诱导的 HSC 继续扩增, 而对照细胞经历细胞死亡或分化。感染后 16 天, 对照的 GFP 转导细胞因分化和死亡而从细胞培养物中耗尽。相比之下, 在过表达 SALL4A 或 SALL4B 的 HSC 中, 细胞继续扩增和存活。在整个细胞培养中, 可见多个健康的正在扩增的细胞簇。

[0037] 图 7. (A) 用 SALL4A、SALL4B 或 GFP 转导并在含有不到 75% 的细胞因子的培养基中培养的 CD34⁺ 细胞的生长曲线。转导后, 在严格条件下培养每组 50,000 个细胞, 在所述严格条件下, 正常细胞因子浓度降低 75%。用 SALL4A 或 SALL4B 转导的 HSC 历经 7 天仍继续存活并扩增, 而对照细胞在第 5 天时停止生长。(B) 在慢病毒-SALL4A 或 -SALL4B 感染后 14 天 CD34⁺/CD38⁻ 细胞相比对照物的扩增倍数。用 SALL4A 转导的细胞呈现 CD34⁺/CD38⁻ 细胞相比对照物的 368 倍增加, 而用 SALL4B 转导的细胞显示 384 倍增加。(C) 慢病毒感染后 31 天, SALL4 诱导的造血干细胞的表型分析。利用人特异性抗体 CD34-PE 和 CD38-APC 来比较 SALL4 转导的 HSC 与 3 天的对照细胞。慢病毒感染后 31 天, 变老的 SALL4 诱导细胞对于 CD34⁺/CD38⁻ 继续呈现与对照细胞类似的表型比率。对三个独立样品进行 FLOW 分析。因此, 这些变老的细胞中有许多仍获得祖细胞特征并且能够分化成各种细胞系。(D) 31 天大的 SALL4 诱导的 HSC 获得类母细胞的形态。31 天大的 SALL4 诱导的 HSC 进行瑞氏-吉姆萨 (Wright-Giemsa) 染色。许多细胞显示出类母细胞的形态, 包括大核和稀少的胞浆。这些细胞代表在 SALL4-慢病毒感染后 31 天仍可见并扩增的未分化细胞群。

[0038] 图 8. 慢病毒感染后 14 天, SALL4 诱导的造血干细胞的代表性表型分析 (n=4)。利用人特异性抗体 CD34-PE 和 CD38-APC 来比较 SALL4 转导的 HSC 与 3 天的对照细胞。

[0039] 图 9. 单个用 SALL4 诱导的 GFP 阳性簇可以解离并扩增形成新的 HSC 簇。从初始细胞培养物中拣取单个过表达 SALL4 的 7 天大的细胞簇, 小心地解离, 并再接种到 48 孔板的单一孔中。4 天后, HSC 能够扩增并形成新的正在增殖的健康细胞簇, 从而使克隆数量净增加。

[0040] 图 10. 18 天大的 SALL4 诱导的 HSC 获得类母细胞的形态。将 18 天大的 SALL4 诱导的 HSC 放入细胞离心机中并进行吉姆萨染色。许多细胞显示出类母细胞的形态, 包括大核和稀少的胞浆。这些细胞代表在 SALL4-慢病毒感染后 18 天仍可见并扩增的未分化细胞群。

[0041] 图 11. CFU 祖细胞是 GFP 阳性的。选择进行 CFU 分析的 HSC 细胞是 GFP 阳性的, 经验证, 这些细胞经历成功转导并过表达 SALL4。

[0042] 图 12. 各种 CFU 集落能够从 SALL4 诱导的 HSC 分化。将变老的用 SALL4 诱导的 HSC 接种于 Methocult 中, 并观察 CFU 集落。在利用 SALL4 诱导的 HSC 进行的 CFU 分析中观察到许多谱系, 包括 BFU-E、CFU-GM 和 CFU-GEMM 集落。这些数据证实, 变老的用 SALL4 转导的 HSC 能够分化成不同血液细胞谱系。

[0043] 图 13. 当生长因子浓度降低 50% 时, SALL4 诱导的 HSC 能够扩增。当 SALL4 转导的 HSC 在含有不到 50% 的细胞因子的生长培养基中培养时, 它们在慢病毒感染后 6 天仍能够存活并扩增。此外, 当生长因子浓度降低到初始值的 25% 时, SALL4 转导的 HSC 继续增殖。

相比之下,对照细胞在第 6 天时经历细胞死亡。

[0044] 图 14. 当生长因子减少 25% 时, SALL4 诱导的 HSC 成功地增殖。当生长因子浓度降低 50% 时, SALL4 诱导的 HSC 能够扩增。当 Sa114 转导的 HSC 在含有不到 50% 的细胞因子的生长培养基中培养时,它们在慢病毒感染后 6 天仍能够存活并扩增。此外,当生长因子浓度降低到初始值的 25% 时, SALL4 转导的 HSC 继续增殖。相比之下,对照细胞在第 6 天时经历细胞死亡。

[0045] 图 15. SALL4 介导的离体 HSC 扩增模型。将原代培养物分开,并用 SALL4 或 GFP 对照物转导。不存在 SALL4 过表达的存活 HSC 的数量因分化或死亡而减少,导致净 HSC 减少。相比之下,对于过表达 SALL4 的 HSC,许多克隆能够在培养中存活并扩增。在整个培养中,展现净 HSC 扩增,具有许多正在扩增的细胞簇。

[0046] 图 16. 慢病毒感染后 14 天和 23 天, SALL4 诱导的造血干细胞的表型分析。利用人特异性抗体 CD34-PE 和 CD38-APC 来比较 SALL4 转导的 HSC 与对照细胞。慢病毒感染后 14 天(顶图)和 23 天(底图),变老的用 SALL4 诱导细胞对于 CD34+/CD38- 继续呈现与对照细胞类似的表型比率。因此,这些变老的细胞中有许多仍获得祖细胞特征并且能够分化成各种细胞系。

[0047] 图 17. SALL4 结合于重要的 HSC 信号传导基因。

[0048] 图 18. 从来自 3 位不同患者的外周血干细胞分离的 CD34+ 细胞。CD34+ 细胞是使用磁性抗 CD34+ 人类微珠从干细胞池分离。富含 CD34+ 的细胞用 SALL4A 转导,并在明视场和荧光显微镜下成像。来自各个患者的全部三份样品都用 SALL4A 成功地转导并且在培养中迅速扩增。

[0049] 图 19. 仅利用 G-CSF 培养的用 SALL4 诱导的 32D 细胞的生长曲线。32D 细胞用 SALL4A、SALL4B 或 GFP 慢病毒转导,随后在含有 IL-3 的生长培养基中培养 3 天。第 4 天,从每组中分出 15000 个细胞,并放入含 G-CSF 并且不含 IL-3 的新生长培养基中。每天监测细胞生长,并记录下每组中存活细胞的数量。在用 SALL4A 转导的细胞中,从第 1 天到第 7 天观察到细胞数量的 8 倍增加。用 SALL4B 转导的细胞展现出细胞的 7 倍扩增。相比之下,仅用 GFP 转导并且 WT (无慢病毒感染)的细胞在相同时间段内呈现细胞数量减少,并且到第 5 天时,几乎所有的细胞都经历细胞死亡。

[0050] 图 20. 32D 细胞的瑞士-吉姆萨染色。在单独 IL3 存在下 (A)、在 G-CSF 存在下用 SALL4A 转导 (B) 以及在单独 G-CSF 存在下 (C) 32D 细胞的形态。给予 IL3 或用 SALL4A 转导的细胞仍呈现类母细胞的形态 (A 或 B),而未用 SALL4A 转导并给予 G-CSF 的细胞展现嗜中性粒细胞形态 (C)。

[0051] 图 21. 将用 HA-SALL4A 和 His-SALL4B 或用 HA-SALL4B 和 His-SALL4A 转染的 293 细胞的溶解产物等量混合并用抗 HA 和抗 His 印迹,显示在这些溶解产物中存在这些蛋白质。使用每一溶解产物混合物,用抗 HA 抗体进行下拉 (pull-down) 并使用抗 His 标签的抗体对所得到的沉淀蛋白质进行蛋白质免疫印迹 (Western blotting)。

[0052] 图 22. (A) 各种 CFU 集落能够从 SALL4 诱导的 HSC 分化。将变老的用 SALL4 诱导的 HSC 接种于 Methocult 中,并观察 CFU 集落。在利用 SALL4 诱导的 HSC 进行的 CFU 分析中观察到许多谱系,包括 CFU-GEMM、BFU-E 和 CFU-GM 集落。这些数据证实,变老的用 SALL4 转导的 HSC 能够分化成不同血液细胞谱系。(B) 由 SALL-4 诱导的造血干细胞形成的 CFU 集

落的数量。在 CFU Methocult 培养基中培养 SALL-4 诱导的细胞或 GFP 诱导的细胞后 13 到 18 天, 计出 CFU 集落的数量。显示了来自第 18 天的代表性数据。(C) 在第 18 天时形成的 CFU 集落的类型。

[0053] 图 23. (A) 在细胞培养 1 个月后, 用 SALL4A 或 SALL4B 转导的 CD34⁺ 细胞相对于对照细胞分别具有 1780 倍和 1463 倍增加。值是平均值 + 标准差(s. d.)。 (B) 此外, 在 1 个月, 对于 LTC-IC 的总数, 相比对照物, 用 SALL4 转导的细胞对于 SALL4A 来说显示出 9.32 倍增加并且对于 SALL4B 来说显示出 8.88 倍增加。值是平均值 + 标准差。(C) 总的来说, 用 SALL4A 转导的细胞相对于对照物具有 16776 倍的总 CD34⁺/CD38⁻ 干细胞扩增倍数, 而用 SALL4B 转导的细胞显示出 13320 倍增加。值是平均值 + 标准差。(D) 来自用 SALL4A 或用 SALL4B 转导的 HSC 移植的 NOD/SCID 接受者的外周血的 CD45⁺ 人白细胞注射后 4 周的代表性流式细胞术分析。(E) 因定植的细胞而展现人类细胞多谱系再群体化的小鼠注射后 4 周的代表性流式细胞术曲线。尽管阴性对照动物未显示人类细胞的定植, 但实验动物显示 CD15⁺ 髓样和 CD19⁺ 淋巴样人类细胞定植。(F) 二次和三次骨髓移植 NOD/SCID 小鼠的 Flow 分析。这些动物在二次(2.74%)和三次(3.29%)移植中都对 CD45⁺ 细胞呈阳性。当进一步分析在三次移植中 CD45⁺ 群体的特定谱系时, CD33 髓样细胞和 CD19/CD3 淋巴样细胞测量呈阳性(G)。(H) 在用 20,000 个(SALL4A (■)、SALL4B (▲)或 GFP (◆))或 40,000 个(SALL4A (□)、SALL4B (△)或 GFP (◇)) 初始人 CD34⁺ 细胞移植的 NOD-SCID 小鼠的外周血中人类嵌合体的量。(I) 在 SALL4A (□)、SALL4B (▲)或 GFP (○) 存在下慢病毒转染后注射到 NOD-SCID 小鼠(n=72) 中的 CD34⁺ 骨髓细胞的限制性稀释分析。

[0054] 图 24. SALL4 诱导的 32D 细胞扩增在去除 IL-3 并添加 G-CSF 之后增殖。在去除 IL-3 并添加 G-CSF 到细胞生长培养基中后 3 天, SALL4A 和 SALL4B 诱导的细胞继续扩增, 而 GFP 诱导的细胞展现细胞数量减少。7 天时, SALL4A 诱导的细胞 (i) 和 SALL4B 诱导的细胞 (ii) 继续增殖, 而对照细胞 (iii) 经历细胞死亡。

[0055] 图 25. (A) 人 His-TAT-SALL4B 构建体的示意性表示。此构建体是通过将人 SALL4B cDNA 克隆到 pTAT-pET28b 载体(来自 S. F. 道迪(S. F. Dowdy) 的礼物, 加利福尼亚大学(University of California), 圣地亚哥(San Diego)) 中来产生。(B) 如通过考马斯蓝(Coomassie blue)染色的 SDS-PAGE 检测的 His-TAT-SALL4B 的亲和纯化。CL, 用 0.1mM IPTG 诱导 3 天的澄清溶解产物; FT, 流过; W, 洗涤液; E, 洗出液。由箭头指示的条带进一步通过 LC-MS/MS 鉴别, 如部分 (D) 中所示。(C) 通过使用抗 6×His 标签的小鼠单克隆抗体进行的纯化 His-TAT-SALL4B 的蛋白质免疫印迹。W, 洗涤液; E, 洗出液。(D) 通过 LC-MS/MS 进行的纯化 His-TAT-SALL4B 蛋白的鉴别。部分 (B) 中的纯化 SALL4B 覆盖的序列的表示。通过 LC-MS/MS 鉴别的 His-TAT-SALL4B 的肽呈黑色。在指定条带中序列覆盖百分比为 34% (58 个肽)。

[0056] 图 26. 当用 TAT-SALL4B 蛋白处理时人骨髓 CD34⁺ 细胞以较高速率扩增。(A) 在蛋白质处理 3 天后 CD34⁺ 细胞的明视场图像。(B) 用 TAT-SALL4B 处理的骨髓细胞相比仅用 BSA 处理的对照细胞的增加倍数和总细胞数量 (C)。(D) 与未操作的 CD34⁺ 细胞相比较, 由用 TAT-SALL4B 蛋白处理的造血干细胞形成的 CFU 集落的数量。

[0057] 图 27. 由 SALL-4 诱导的造血干细胞形成的 CFU 集落的数量。在 CFU methocult 培养基中培养 SALL-4 诱导的细胞或 GFP 诱导的细胞之后 13 到 18 天, 计出 CFU 集落的数量和

类型。

[0058] 图 28. 在 NOD/SCID 小鼠中定植的人类细胞的多谱系再群体化。因定植的细胞而展现人类细胞的多谱系再群体化的小鼠骨髓的代表性流式细胞术曲线。注射后 15 周, 实验动物显示 CD3+/CD19+ 淋巴样、CD15+ 髓样和血型糖蛋白 -4A+ 红系人类细胞定植。

[0059] 图 29. 从感染杆状病毒的 Sf9 细胞分离的 His-TAT SALL4B 的纯化的蛋白质免疫印迹。所用抗体是抗 SALL4 (亚诺法公司 (Abnova), 台湾台北市 (Taipei City, Taiwan)), 1:2000 ; 和抗小鼠 HRP [山羊], 亚诺法公司, 1:2000。

[0060] 图 30. 用杆状病毒感染的 Sf9 细胞, 第 4 天。

[0061] 图 31. TAT-SALL4B 增加小鼠全骨髓细胞的增殖。

[0062] 图 32. 在用 SALL4 处理的动物中骨髓恢复并且骨髓细胞扩增 (实验设计)。

[0063] 图 33. 在注射 TAT-SALL4B 后骨髓细胞扩增。

[0064] 图 34. 在小鼠中腹膜注射 PBS、G-CSF 和 SALL4B 之后骨髓细胞扩增的比较。骨髓腔的组织切片显示骨髓细胞扩增。

[0065] 图 35. 来自用 PBS、SALL4B 或 G-CSF 处理的小鼠的骨髓细胞的 CFC 数量。C57B/6 小鼠进行致死性照射 (7Gy), 并且在照射 7 天后接受 24 小时治疗。第 8 天, 分离小鼠骨髓细胞并在 MethoCult 中培养以进行 CFC 分析。每 20,000 个全骨髓细胞中, PBS 组、SALL4B 组和 G-CSF 组的第 7 天 CFC 分别为 6 ± 1.41 、 11.3 ± 2.51 和 18 ± 2.82 (PBS 相比 SALL4B、PBS 相比 G-CSF、SALL4B 相比 G-CSF 的 P 值 < 0.05)。

[0066] 图 36. 来自用 SALL4B、G-CSF 或 PBS 处理的动物的 HSC (Sca1+/c-kit+) 相比对照物 (PBS) 的增加倍数。

[0067] 图 37. 在 NOD-SCID 小鼠的外周血中人 UCB 长期定植的增加倍数。将约 20,000 个脐血源性 CD34+ 细胞移植到经历亚致死照射的 NOD/SCID 小鼠中, 并用 2 微克 / 天 SALL4B 蛋白治疗 7 天, 随后每隔一天用 2 微克 / 天再治疗一周。移植后 16 周 CD45 细胞 (UCB 细胞) 与 PBS 处理的对照物的比较将通过流式细胞术测量。

[0068] 图 38. SALL4B 增进 NOD/SCID 小鼠中人脐血细胞的长期定植。在经历亚致死照射后 24 小时, 将 20,000 个人脐血 CD34+ 细胞移植到 NOD/SCID 小鼠中 (2.5Gy)。小鼠用 2 μ g 的 TAT-SALL4B 蛋白或 PBS 处理 2 周 (第一周每天一次并且第二周每隔一天一次)。移植后 4 个月, 收集小鼠骨髓细胞, 并通过流式细胞术针对 CD45 阳性细胞 (定植的脐血细胞) 进行分析。

[0069] 图 39. SALL4B 增加外周血 (PB) 的 CFC 数量。野生型 C57B/6 小鼠接受 PBS 或 6 μ g 的 TAT-SALL4B 蛋白注射 (腹膜内注射) 达 5 天, 并在最后一次注射后 2 小时收集 PB。使用来自经过注射的小鼠的外周血的成核细胞进行 CFC (集落形成祖细胞) 分析。在显微镜下计出第 7 天 CFC 的数量。

具体实施方式

[0070] 本发明的示范性实施例

[0071] 本发明提供一种用于扩增干细胞群的方法, 所述方法包含向所述干细胞群提供有效扩增所述干细胞群的量的多肽, 所述多肽具有 Sa1 样 (SALL) 多肽的增强扩增活性。

[0072] 在一个实施例中, 干细胞是成体干细胞。

[0073] 在一些实施例中,干细胞是在以下各项中或来源于以下各项:脑、肝、心脏、肾、皮肤、胰腺、膀胱、胆囊、大肠、小肠、胃、骨骼肌或肺。

[0074] 在另一实施例中,干细胞是造血干细胞。

[0075] 在另一实施例中,造血干细胞是在以下各项中或来源于以下各项:脐带血、外周血、骨髓或脾。在另一实施例中,造血干细胞是人类干细胞。

[0076] 在一些实施例中,用于扩增干细胞群的方法也包含授予 SALL 多肽与刺激因子的组合。在一些实施例中,刺激因子是 G-SCF、GM-CSF、M-CSF、干细胞因子或 FMS 样酪氨酸激酶-3 (FMS-like tyrosine kinase-3, FLT-3)。

[0077] 在一些实施例中,干细胞群是离体扩增。

[0078] 在一些实施例中,干细胞群是在体内扩增。在另一实施例中,干细胞群是在包含 50ng/ml 的 FMS 样酪氨酸激酶-3 (FLT-3)、50ng/ml 血小板生成素 (Thrombopoietin, TPO) 和 / 或 50ng/ml 干细胞因子 (Stem cell factor, SCF) 的培养基中培养。

[0079] 在另一实施例中,干细胞群是在包含 25ng/ml 的 FMS 样酪氨酸激酶-3 (FLT-3)、25ng/ml 血小板生成素 (TPO) 和 / 或 25ng/ml 干细胞因子 (SCF) 的培养基中培养。

[0080] 在一个实施例中, SALL 多肽连接到转运部分 (transport moiety), 所述转运部分能够跨过细胞膜, 由此将 SALL 多肽转运到细胞中。

[0081] 在另一实施例中, 转运部分是 HIV-1 转录转活化因子 (transactivator of transcription, TAT) 肽、Chariot 蛋白、富精氨酸肽、触足源性穿膜肽、1 型单纯疱疹病毒 VP22 蛋白或 +36GFP。

[0082] 在另一实施例中, 用于本发明中的 SALL 可以呈核酸或多肽的形式。

[0083] 在一个实施例中, SALL 多肽包含 SEQ ID No: 7、8、9、10、11 或 12 中所陈述的氨基酸序列。

[0084] 在一个实施例中, SALL 多肽是由包含 SEQ ID No: 1、2、3、4、5 或 6 的核苷酸序列编码。

[0085] 在另一实施例中, 所述干细胞群具有 SALL 多肽, 所述 SALL 多肽包含 SEQ ID No: 7、8、9、10、11 和 / 或 12 中所陈述的氨基酸序列。

[0086] 在一个实施例中, 所述群体中的细胞用包含编码 SALL 多肽的核苷酸的病毒载体转导, 由此将 SALL 多肽提供给干细胞群。

[0087] 在另一实施例中, 所述病毒载体是来源于 (但不限于) 腺病毒、杆状病毒、细小病毒、疱疹病毒、痘病毒、腺相关病毒、塞姆利基森林病毒 (Semliki Forest virus)、痘苗病毒和逆转录病毒。

[0088] 在一个实施例中, 在干细胞中瞬时表达外源 SALL 基因。

[0089] 在另一实施例中, 编码 SALL 多肽的核苷酸包含如 SEQ ID No: 1、2、3、4、5 或 6 中所陈述的序列。

[0090] 在另一实施例中, 编码 SALL 多肽的核苷酸是在诱导型启动子的控制下表达。

[0091] 在另一实施例中, 用于本发明中的 SALL 多肽可以是外源或内源的。优选 SALL 多肽是外源的。在本发明中, 将外源 SALL 多肽提供给细胞, 由此促进细胞的扩增。在这种情况下, 内源 SALL 多肽可以补充有外源 SALL 多肽以增进其作用。

[0092] 在一个实施例中, 干细胞群扩增 10 倍、20 倍、50 倍、100 倍或 1000 倍。

[0093] 本发明另外提供一种用于增进受检者体内干细胞群的扩增的组合物,所述组合物包含有效扩增所述干细胞群的量的 SALL 多肽,和培养基。在一个实施例中,所述组合物另外包含干细胞。

[0094] 本发明另外提供一种鉴别用于扩增干细胞群的试剂的方法,所述方法包含 (a) 获得候选试剂;(b) 将来自所述群体的干细胞暴露于所述候选试剂;和 (c) 确定在所述干细胞中 SALL 多肽是否上调,其中如果 SALL 多肽上调,那么确定所述试剂是用于扩增所述干细胞群的试剂。在一个实施例中,候选试剂可以提供于文库中。在另一实施例中,本发明提供一种用于扩增通过上述筛选方法获得的干细胞的试剂。

[0095] 本发明另外提供一种用于治疗或预防需要干细胞或由其得到的扩增的干细胞的受检者的疾病、病症或异常的方法,所述方法包含 a) 获得干细胞群;b) 向所述干细胞群提供有效扩增所述干细胞群的量的 SALL 多肽;和 c) 将有效治疗或预防所述受检者的所述疾病、病症或异常的量的扩增的干细胞群移植到所述受检者。

[0096] 在一个实施例中,本发明可以靶向造血和循环(血液细胞等)疾病、病症或异常,包括(但不限于)贫血(例如再生障碍性贫血(特别是严重再生障碍性贫血)、肾性贫血、癌性贫血、继发性贫血、难治性贫血等)、癌症或肿瘤(例如白血病);以及在其化疗后,造血功能衰竭、血小板减少症、急性髓细胞白血病(特别是首次缓解(高危群体)、二次缓解和其后的缓解)、急性淋巴细胞白血病(特别是首次缓解、二次缓解和其后的缓解)、慢性髓细胞白血病(特别是慢性期、反式迁移期(transmigration period))、恶性淋巴瘤(特别是首次缓解(高危群体)、二次缓解和其后的缓解)、多发性骨髓瘤(特别是发作后的早期)等。本发明还靶向心脏衰竭、心绞痛、心肌梗塞、心律不齐、瓣膜性心脏病、心肌/心包疾病、先天性心脏病(例如房间隔缺损、室间隔缺损、动脉导管未闭、法洛四联症(tetralogy of Fallot))、动脉疾病(例如动脉硬化、动脉瘤等)、静脉疾病(例如静脉曲张等)和淋巴管疾病(例如淋巴水肿)、镰状细胞疾病以及放射诱发损伤的治疗、自体免疫疾病、脑性瘫痪、重症肢体缺血、退化性关节疾病、2 型糖尿病、心脏衰竭、多发性硬化、骨关节炎、类风湿性关节炎和脊髓损伤。

[0097] 本发明另外提供一种干细胞库,包含遗传上不同的干细胞群,其中所述干细胞群已经根据本发明的方法扩增。

[0098] 本发明提供一种用于扩增受检者体内的干细胞群的方法,包含增加所述受检者体内的 SALL 多肽量。

[0099] 本发明另外提供一种用于增加受检者体内祖细胞群的规模的方法,所述方法包含增加所述受检者体内的 SALL 多肽量。

[0100] 在一些实施例中,增加所述受检者体内的 SALL 多肽量包含将所述 SALL 多肽引入所述受检者体内。

[0101] 在一些实施例中,增加所述受检者体内的 SALL 多肽量包含将核酸分子引入所述受检者体内,所述核酸分子具有编码 SALL 多肽的核苷酸序列。

[0102] 在一些实施例中,编码 SALL 多肽的核苷酸包含如 SEQ ID No:1、2、3、4、5 或 6 中所陈述的核苷酸序列。

[0103] 在一些实施例中,SALL 多肽包含转运部分,所述转运部分能够跨过细胞膜,由此将 SALL 多肽转运到细胞中。

[0104] 在一些实施例中,祖细胞是造血祖细胞群。

[0105] 本发明多个方面还提供增进受检者体内造血干细胞的长期定植的方法。在一些实施例中,造血干细胞是来源于供体。

[0106] 需要干细胞或由其得到的扩增的干细胞的受检者的疾病、病症或异常的非限制性实例包括(但不限于)严重再生障碍性贫血、白细胞减少症、嗜中性粒细胞减少症、急性放射综合征、多发性骨髓瘤、非霍奇金氏淋巴瘤(non-Hodgkin's lymphoma)和霍奇金氏淋巴瘤。本发明的方法可以与已知会产生例如白细胞数量减少等副作用的各种癌症治疗(例如化疗、放射疗法和骨髓移植)组合使用。

[0107] 预期本文中公开的每一实施例适用于每一其它公开的实施例。因此,本文中描述的各种要素的所有组合都在本发明的范围内。

[0108] 应了解,在提供参数范围的情况下,本发明也提供在此范围内的所有整数和其十分位数。举例来说,“每天 0.2-5mg/kg”包括每天 0.2mg/kg、每天 0.3mg/kg、每天 0.4mg/kg、每天 0.5mg/kg、每天 0.6mg/kg 等,最多、每天 5.0mg/kg。

[0109] 定义

[0110] 如本文中所用,干细胞是指能够无限地自更新并且分化产生至少一种功能性、终末细胞类型的细胞。

[0111] 如本文中所用,祖细胞是指自更新能力有限并且分化产生至少一种功能性、终末细胞类型的细胞。

[0112] 如本文中所用,造血干细胞是指通常在血液中发现的细胞以及产生通常在血液中发现的细胞(例如在骨髓中发现的细胞)的细胞。在此情形中“通常”包括一个人经历治疗以改变血液或骨髓中细胞的数量或质量的情况。

[0113] 如本文中所用,造血干细胞是指产生所有血液细胞类型的专能干细胞。

[0114] 如本文中所用,“扩增”是指通过使干细胞增殖来增加干细胞的数量,与将非干细胞的细胞转变成干细胞相对。

[0115] 在本文中使用的 Sal 样(SALL)基因或多肽包含 SALL1、SALL2、SALL3 和 SALL4,它们最初是基于与果蝇基因 sal 同源的 DNA 序列克隆。SALL4 包含 SALL4a 和 SALL4b。

[0116] 病毒载体在本文中用于指包含能够被引入细胞中并表达的全部或部分病毒基因组的载体。这些病毒载体可以包括原生病毒、突变病毒或重组病毒。这些病毒可以具有 RNA 或 DNA 基因组。适合病毒载体的实例包括(但不限于)腺病毒、杆状病毒、细小病毒、疱疹病毒、痘病毒、腺相关病毒、塞姆利基森林病毒、痘苗病毒以及逆转录病毒和杂合载体。

[0117] “转运部分”在本文中用于指能够跨过细胞膜并且可以将本发明的多肽转运到干细胞中的多肽。转运部分的实例包括(但不限于)HIV-1 转录转活化因子(Tat)肽、ChariotTM 蛋白、富精氨酸肽、触足源性穿膜肽、1 型单纯疱疹病毒 VP22 蛋白和 +36GFP。在 SALL 多肽连接到转运部分的实施例中,“连接”可以是共价连接。

[0118] 术语“有效量”意指将引起研究人员、兽医、药师或其它临床医师所寻求的细胞、组织、系统、动物或人类的生物学或药物反应的本发明多肽或干细胞的量。

[0119] “构建体”用于指产生的用以达到表达特定核苷酸序列的目的或打算用于构建其它重组核酸的重组核酸,所述重组核酸可以是重组 DNA 或 RNA 分子。一般说来,“构建体”在本文中用于指分离的重组 DNA 或 RNA 分子。

[0120] 术语“能够定植”在此用于指造血细胞植入骨髓中保持较长时间段(例如至少一

年)的能力。植入可以直接检测、(例如通过活组织检查)或通过在血液中子代细胞的产生来检测。

[0121] 转导用于指通过使用病毒载体将遗传物质引入细胞中。如本文中所用,转导的细胞由转导过程产生,并且含有在转导过程之前其不含的遗传物质,不管是否为稳定整合的。

[0122] 短语“培养基”用于指用于培养干细胞的任何标准培养基。

[0123] 如本文中所用,“核酸序列”是指寡核苷酸或聚核苷酸及其片段或部分,以及基因组或合成来源的DNA或RNA,它可以是单链或双链的,并且代表有义链或反义链。类似地,如本文中所用,“氨基酸序列”是指寡肽、肽、多肽或蛋白质序列,及其片段或部分,以及天然存在或合成的分子。

[0124] 在本说明书通篇,“包含”一词和其变化形式应理解为暗示包括所述要素、整数或步骤,或者要素、整数或步骤组在内,但不排除任何其它要素、整数或步骤,或者要素、整数或步骤组。

[0125] 干细胞

[0126] 本发明的干细胞(例如成体干细胞和造血干细胞)包括已经在哺乳动物器官或组织中鉴别出来的本领域中已知的所有干细胞。干细胞可以包括(但不限于)多能干细胞、间充质干细胞、造血干细胞、髓样干细胞和淋巴样干细胞。得到最佳表征的是造血干细胞。造血干细胞是从骨髓、血液、脐血、胎肝和卵黄囊分离,是产生血液细胞或在移植后再起始多个造血谱系并且可以再起始造血作用延续接受者的生命的祖细胞。(参看,费R.(Fei, R.)等人,美国专利第5,635,387号;麦克格拉夫(McGlave)等人,美国专利第5,460,964号;西蒙斯P.(Simmons, P.)等人,美国专利第5,677,136号;塚本(Tsukamoto)等人,美国专利第5,750,397号;斯查沃兹(Schwartz)等人,美国专利第5,759,793号;迪奎斯妥(DiGuisto)等人,美国专利第5,681,599号;塚本等人,美国专利第5,716,827号;希尔B.(Hill, B.)等人,1996。)当移植到经过致死照射的动物或人体中时,造血干细胞可以使红系细胞、嗜中性粒细胞-巨噬细胞、巨核细胞和淋巴样造血细胞池再群体化。

[0127] 本领域中众所周知,造血细胞包括专能干细胞(例如淋巴样干细胞)和/或提供给特定造血谱系的祖细胞。提供给特定造血谱系的祖细胞可以属于T细胞谱系、B细胞谱系、树突状细胞谱系、朗格汉斯细胞(Langerhans cell)谱系和/或淋巴样组织特异性巨噬细胞谱系。造血干细胞和祖细胞谱系论述于夏伯格(Sieburg)等人,2006;斯查罗德(Schroeder)等人,2010;戴卡斯特拉(Dykstra)等人,2007;以及美国专利第7,994,114号中,各案的内容都以引用的方式并入本文中。

[0128] 造血干细胞可以从血液制品获得。如本发明中所用,“血液制品”定义为从身体或含有造血系统来源的细胞的身体器官获得的产品。这些来源包括未分级骨髓、脐带、外周血、肝、胸腺、淋巴和脾。本领域技术人员将显而易见,所有上文提到的粗制或未分级血液制品都可以通过多种方式而富集具有“造血干细胞”特征的细胞。举例来说,血液制品可以从分化程度更高的子代耗尽。更成熟、分化程度更高的细胞可以针对、经由其表达的细胞表面分子进行选择。另外,可以分级分离血液制品,并针对CD34⁺细胞进行选择。CD34⁺细胞在本领域中被认为包括能够自更新并具有多能性的细胞亚群。所述选择可以使用例如可商购的抗CD34磁珠(戴诺公司(Dynal),纽约州成功湖(Lake Success, N.Y.))实现。未分级的血液制品可以直接从供体获得或从低温保存的储藏重新得到。

[0129] 造血干细胞可以在离体扩增干细胞群之前采集或收集。“采集”造血祖细胞定义为从基质移取或分离细胞。这可以使用多种方法实现,例如酶法、非酶法、离心法、电学方法或基于尺寸的方法,或优选通过使用培养基(例如培育细胞的培养基)冲洗细胞。细胞可以进一步收集、分离,并且再使用本发明扩增,并产生较大的群体。

[0130] 用于分离造血干细胞的方法是本领域中众所周知的,并且通常涉及基于细胞表面标记物和功能特征的后续纯化技术。造血干细胞和祖细胞可以从骨髓、血液、脐血、胎肝和卵黄囊分离,并产生多个造血细胞谱系,并且可以再起始造血作用延续接受者的生命。(参看,费 R. 等人,美国专利第 5,635,387 号;麦克格拉夫等人,美国专利第 5,460,964 号;西蒙斯等人,美国专利第 5,677,136 号;塚本等人,美国专利第 5,750,397 号;斯查沃兹等人,美国专利第 5,759,793 号;迪奎斯妥等人,美国专利第 5,681,599 号;塚本等人,美国专利第 5,716,827 号;希尔 B. 等人,1996。)举例来说,为了从外周血分离造血干细胞和祖细胞,将含血液的 PBS 装入 Ficoll 管(Ficoll-Paque,安玛西亚公司(Amersham))中,并以 1500rpm 离心 25 到 30 分钟。离心后,收集白色的中央环作为含有造血干细胞。

[0131] 干细胞可以从例如脑、肝、心脏、肾、皮肤和肺等其它组织分离。

[0132] 干细胞移植

[0133] 干细胞移植可以用作例如白血病、某些类型的淋巴瘤(包括霍奇金淋巴瘤)、再生障碍性贫血、地中海贫血、镰状细胞贫血等血液病症以及一些先天性代谢或免疫缺陷障碍(例如慢性肉芽肿病)的治疗的一部分。某些类型的干细胞也可以用作移植用于骨髓被治疗一些癌症所使用的高剂量化疗或放射疗法破坏的患者。

[0134] 干细胞可以是受检者自身的细胞(自体移植)或供体的细胞(异基因移植)。当使用受检者自身的干细胞时,它们可以在化疗或放射疗法之前收集,即“体内收集”,因为这些治疗会破坏干细胞。它们可以在处理(例如增加或扩增(“体内扩增”)所扩增的细胞的质量)后回输到体内。

[0135] 当前用于干细胞策略的方法涉及从胚胎脐血收集干细胞和由血液捐献收集干细胞。这些技术的应用有限,因为外周血中干细胞总数占较小比例,并且因为只能从输血收集到有限量的血液。使用来自成体的干细胞的一个优势在于,可以使用本发明中描述的方法在培养中扩增受检者自身的细胞,随后再引入受检者体内。因此,在收集人干细胞群以供长期低温储存(例如在干细胞库中),以及用于最终解冻低温保存的细胞群以治疗由自体转移引起的疾病方面也存在未满足的需求。供低温储存的干细胞的保存是本领域中众所周知的。举例来说,参看,人类干细胞的培养(Culture of human stem cells)(威立-利斯公司(Wiley-Liss),2007),和低温保存和冷冻干燥方案(Cryopreservation and freeze-drying protocols)(胡马纳出版社(Humana Press),2007)。这些文献以引用的方式并入本文中。

[0136] 本发明扩增的干细胞群可以用于治疗造血和循环(血液细胞等)疾病、病症或异常。所述疾病、病症或异常的实例包括(但不限于)贫血(例如再生障碍性贫血(特别是严重再生障碍性贫血)、肾性贫血、癌性贫血、继发性贫血、难治性贫血等)、癌症或肿瘤(例如白血病);以及在其化疗后,造血功能衰竭、血小板减少症、急性髓细胞白血病(特别是首次缓解(高危群体)、二次缓解和其后的缓解)、急性淋巴细胞白血病(特别是首次缓解、二次缓解和其后的缓解)、慢性髓细胞白血病(特别是慢性期、反式迁移期)、恶性淋巴瘤(特别是首次

缓解(高危群体)、二次缓解和其后的缓解)、多发性骨髓瘤(特别是发作后的早期)等。本发明还靶向心脏衰竭、心绞痛、心肌梗塞、心律不齐、瓣膜性心脏病、心肌/心包疾病、先天性心脏病(例如房间隔缺损、室间隔缺损、动脉导管未闭、法洛四联症)、动脉疾病(例如动脉硬化、动脉瘤等)、静脉疾病(例如静脉曲张等)和淋巴管疾病(例如淋巴水肿)、镰状细胞疾病以及放射诱发损伤的治疗、自体免疫疾病、脑性瘫痪、重症肢体缺血、退化性关节疾病、2型糖尿病、心脏衰竭、多发性硬化、骨关节炎、类风湿性关节炎和脊髓损伤。

[0137] 有关干细胞移植以治疗人类疾病、病症或异常的方法是本领域中众所周知的。举例来说,参看,干细胞和骨髓移植手册(Manual of Stem Cell and Bone Marrow Transplantation)(剑桥大学出版社(Cambridge University Press),2009)、干细胞移植:生物学、加工和疗法(Stem cell transplantation:biology,processing, and therapy)(威立-VCH公司(Wiley-VCH),2006),以及造血干细胞移植实践(Practical Hematopoietic Stem Cell Transplantation)(威立-布拉克维尔公司(Wiley-Blackwell),2007)。这些文献以引用的方式并入本文中。

[0138] 干细胞定植

[0139] 在移植后骨髓再生随移植的细胞的适当定植而变化。在优选实施例中,移植细胞的定植是细胞的长期定植。在一些实施例中,本发明涵盖来源于人脐带血的造血干细胞(HSC)的定植改善。本发明的实施例可用于治疗各种疾病,例如骨髓衰竭病症、各种遗传疾病和造血系统恶性肿瘤。

[0140] 在将来自供体的造血干细胞(HSC)移植到宿主体内的情形中,本发明的实施例可用于增加供体细胞与宿主细胞的嵌合。如本文中所用,嵌合是两种遗传上不同类型的细胞共存于单一生物体中。一旦确立了嵌合,HSC就可以在宿主内增殖。

[0141] 如本文中所公开,腹膜内注射的TAT-SALL4蛋白急剧增加嵌合性和干细胞长期定植。根据本发明实施例,本文所述的SALL多肽可以用于显著增加移植的细胞在骨髓或小生境中的存活率、扩增和定植或嵌合。

[0142] 本发明多个方面涉及基因疗法,特别是施用于造血干细胞和造血祖细胞,转导的细胞,以及获得这些细胞的方法,和使用其提供经过修饰的造血细胞在受检者体内的长期定植的方法。

[0143] 多肽

[0144] 本发明涵盖一种具有Sal样多肽的增强扩增活性的多肽。术语“具有Sal样多肽的增强扩增活性的多肽”是指如在特定生物分析中所测量,展现类似于但未必等同于本发明多肽的活性的多肽。

[0145] 本领域技术人员将显而易见,使SALL的氨基酸序列改变的DNA序列多形性可以存在于群体(例如人群)内。SALL基因中的这种遗传多形性可以因自然等位基因变异而存在于群体内的个体之间。如本文中所用,术语“基因”和“重组基因”是指包含编码SALL多肽(优选哺乳动物SALL多肽)的开放阅读框的核酸分子。这些自然等位基因变异通常会引起SALL基因的核苷酸序列的1%到5%的变化。SALL中由自然等位基因变异引起并且不会改变SALL的功能活性的任何和所有这些核苷酸变异以及由此引起的氨基酸多形性预期在本发明的范围内。本发明涵盖与本文中所公开的SALL核苷酸实质上同源的SALL多肽中的核苷酸变异和由此引起的氨基酸多形性。

[0146] 为了测定两个氨基酸序列或两个核酸的同源性百分比,可以比对这些序列来达到最佳比较的目的(例如,可以在一种蛋白质或核酸的序列中引入空位,以便与另一蛋白质或核酸最佳比对)。随后比较在对应氨基酸位置或核苷酸位置处的氨基酸残基或核苷酸。当占据一个序列中某一位置的氨基酸残基或核苷酸与另一序列中的对应位置相同时,则所述分子在该位置同源。如本文中所用,氨基酸或核酸“同源性”相当于氨基酸或核酸“一致性”。两个序列之间的同源性百分比随这些序列所共有的一致位置的数量而变化(即,同源性 % = 一致位置的数量 / 位置的总数 × 100)。

[0147] 本发明公开了在 HSC 中过表达 SALL4a 和 SALL4b 多肽的结果,并且证实,这两种多肽的过表达将增进 HSC 的离体扩增。SALL4a 和 SALL4b 是 SALL 基因的异构体,并且分别为 1053 个和 616 个氨基酸长。然而,即使 SALL4b 缺少 SALL4a 中存在的超过 400 个氨基酸,但它仍能够增进如本文所公开的 HSC 的扩增,达到 SALL4a 的水平。因此,SALL4a 中并非所有氨基酸都是 Sa1 样多肽的增强扩增活性所需的。根据 ALIGN Query(GENESTREAM SEARCH 网络服务器 IGH Montpellier, 法国,) SALL4a 与 SALL4b 的总体序列一致性为 57.5%。然而,使用 ExPASy (专业蛋白质分析系统(Expert Protein Analysis System))蛋白质组学服务器进行 SALL4a 和 SALL4b 多肽的序列比对揭露,SALL4a 中的两个结构域可能对于这些多肽在 HSC 中的生物活性极为重要。SALL4a 和 SALL4b 的假定结构域 1 包括 SALL4a 和 SALL4b 的前 385 个 N 末端氨基酸,并且揭露这两个多肽之间存在 98.7% 序列一致性。假定结构域 2 包括 SALL4a 的 C 末端氨基酸 822 到 1053,以及 SALL4b 的 C 末端氨基酸 385 到 616,并且揭露这两个多肽之间存在 97% 序列一致性。根据这一信息,有可能容易地制备出具有 SALL4a 或 SALL4b 的增强活性的其它肽。

[0148] 本发明也涵盖这样一种多肽,这些多肽与具有 Sa1 样多肽的增强扩增活性的多肽具有较低程度一致性,但具有足够相似性,由此执行与 SALL 多肽所执行相同的一种或一种以上功能。相似性是由保守氨基酸取代来测定。所述取代是利用特征类似的另一氨基酸取代多肽中的指定氨基酸的取代。保守性取代可能是表型沉默的。通常被视为保守性取代的有:脂肪族氨基酸 Ala、Val、Leu 和 Ile 间的相互置换;羟基残基 Ser 与 Thr 的相互交换;酸性残基 Asp 与 Glu 的交换;酰胺残基 Asn 与 Gln 之间的取代;碱性残基 Lys 与 Arg 的交换;以及芳香族残基 Phe、Tyr 间的置换。有关何种氨基酸改变可能是表型沉默的指导见于鲍维(Bowie)等人(1990)科学(Science)247:1306-1310 中。

[0149] 根据 ALIGNQuery, SALL1 转录变异体 1、SALL1 转录变异体 2、SALL2 和 SALL3 中每一者与 SALL4a 的总体序列一致性分别为 38.1%、37.1%、28.1% 和 38.4%。

[0150] 如本文中所用,多肽(或所述多肽的一个区域)与具有 Sa1 样多肽的增强扩增活性的多肽实质上同源,此时氨基酸序列至少约 25-30%、30-35%、35-40%、40-45%、45-50%、50-55%、55-60%,通常至少约 70-75%,更通常至少约 80-85%,并且最通常至少约 90-95% 或 95% 以上同源。根据本发明,实质上同源的氨基酸序列将由与具有 SEQ ID NO:1、2、3、4、5 和 / 或 6 中所示序列的核酸序列或其部分在严格条件下杂交的核酸序列编码。

[0151] 本发明另外涵盖因遗传密码简并性而与所公开的编码 SALL 多肽的核苷酸序列不同的核酸分子。因此,这些核酸编码与 SEQ ID NO:1、2、3、4、5 和 / 6 中所示核苷酸序列所编码的多肽相同的 Sa1 样多肽。

[0152] 编码“具有 Sa1 样多肽的增强扩增活性的多肽”的核酸片段可以通过以下方式制

备:分离 SEQ ID NO:1、2、3、4、5 和 / 或 6 的一部分,所述序列编码具有 Sal 样多肽的生物活性的多肽;表达 Sal 样多肽的编码部分(例如,通过离体重组表达);并评估 Sal 样多肽的编码部分的活性。

[0153] 增加受检者体内 SALL 多肽的量

[0154] 可以通过多种方式增加受检者体内 SALL 多肽的量,包括增加 SALL 多肽的表达的试剂、授予 SALL 多肽以及在受检者体内引入编码 SALL 肽的核苷酸。

[0155] 可以用于增加 SALL 肽的表达的试剂包括(但不限于)肽、肽模拟物、寡核苷酸、小分子化合物和诱导 RNA 干扰的分子。

[0156] 对受检者授予 SALL 多肽可以经鼻、肺、不经肠、静脉内(i.v.)、腹膜内(i.p.)、关节内、透皮、皮内、皮下(s.c.)、局部、肌肉内、直肠、鞘内、眼内和口腔授予,都使用药学领域中的普通技术人员众所周知的剂型进行。优选的 SALL 多肽授予途径包括静脉内和腹膜内。

[0157] SALL 多肽可以与针对预定投药形式适当选择并且与常规药学实践相符的适合药物稀释剂或载剂(在本文中统称为药学上可接受的载剂)混合授予。投药单元可以是适合局部、静脉内或直接注射或者不经肠授予的形式。

[0158] 用于制备适用于本发明中的剂型的一般技术和组合物描述于以下参考文献中:7 现代药剂学(Modern Pharmaceutics),第 9 和 10 章(班克(Banker)和罗德(Rhodes)编,1979);药物剂型:片剂(Pharmaceutical Dosage Forms:Tablets)(利伯曼(Lieberman)等人,1981);安塞尔(Ansel),药物剂型介绍(Introduction to Pharmaceutical Dosage Forms)第 2 版(1976);雷氏药学大全(Remington's Pharmaceutical Sciences),第 17 版(默克出版公司(Mack Publishing Company),宾夕法尼亚州伊斯顿(Easton, Pa.),1985);现代制药学(Advances in Pharmaceutical Sciences)(大卫盖德顿(David Ganderton),泰文琼斯(Trevor Jones)编,1992);现代制药学,第 7 卷.(大卫盖德顿,泰文琼斯,詹姆斯麦克吉提(James Mc Ginity)编,1995);用于药物剂型的水性聚合物涂层(Aqueous Polymeric Coatings for Pharmaceutical Dosage Forms)(药物与制药学(Drugs and the Pharmaceutical Sciences),第 36 辑(詹姆斯麦克吉提编,1989);药用颗粒载剂:治疗应用:药物和制药学(Pharmaceutical Particulate Carriers:Therapeutic Applications:Drugs and the Pharmaceutical Sciences),第 61 卷(艾伦罗兰德(Alain Rolland)编,1993);胃肠道药物递送(Drug Delivery to the Gastrointestinal Tract)(艾利斯霍德生物学丛书(Ellis Horwood Books in the Biological Sciences). 制药技术系列(Series in Pharmaceutical Technology);J.G. 哈迪(J.G.Hardy),S.S. 戴维斯(S.S.Davis),克里夫 G. 韦尔森(Clive G.Wilson)编);现代药物和制药学(Modern Pharmaceutics Drugs and the Pharmaceutical Sciences),第 40 卷(吉尔伯特 S. 班克(Gilbert S.Banker),克里斯托弗 T. 罗德(Christopher T.Rhodes)编)。这些参考文献的全文以引用的方式并入本申请案中。

[0159] 本发明还提供通过在受检者体内引入编码 SALL 多肽的核苷酸来增加受检者体内 SALL 多肽的量。在一些实施例中,编码 SALL 多肽的核苷酸包含如 SEQ ID No:1、2、3、4、5 或 6 中所陈述的核苷酸序列。编码 SALL 多肽的核苷酸可以可操作地连接到例如启动子等表达调控序列,并且可以通过使用如下文所述的适合载体引入受检者体内。

[0160] 载体

[0161] 可以使用不同类型的载体来转导或转化干细胞。这些包括质粒或病毒载体。到目前为止,逆转录病毒载体已经被广泛用于基因疗法中,特别是基于莫洛尼鼠类白血病病毒(Moloney murine leukemia virus,MoMLV)的载体。基于鼠类逆转录病毒的载体可以用于高效地转导细胞,然而,它们需要细胞在细胞分裂时具有活性。因此,用基于鼠类逆转录病毒的载体转导 HP 细胞需要细胞的活化。

[0162] 慢病毒载体(逆转录病毒载体的一个子类)也可以用于高效地转导(哈斯(Haas)等人,2000;三吉(Miyoshi)等人,1999;卡斯(Case)等人,1999),并且能够转导非分裂细胞,无需诱导 HP 细胞进入细胞周期。此举将避免在一些细胞中的细胞周期诱导可能引起的多能性丧失。其它类逆转录病毒,例如泡沫病毒属(例如泡沫病毒)(瓦斯洛普罗斯(Vassilopoulos)等人,2001)也能够有效地转导非分裂细胞。

[0163] 可以用于本发明中的其它类型的病毒载体包括腺病毒载体(范(Fan)等人,2000;

[0164] 柯南-山泽尔(Knaan-Shanzer)等人,2001;马里尼(Marini)等人,2000)、腺相关病毒(adeno-associated viral,AAV)载体(费舍尔-亚当斯(Fisher-Adams)等人,1996)、基于 SV40 的载体(斯特埃尔(Strayer)等人,2000)或杂合载体形式(例如冯(Feng)等人,1997)。腺病毒载体可以容易地以高滴度产生,并且可以转导非分裂细胞。

[0165] AAV 载体是不致病的,转导增殖和非增殖细胞(包括 CD34+ 细胞),并且稳定地整合到细胞基因组中(格里姆(Grimm)和克莱恩斯查米特(Kleinschmidt)1999)。此外,它们不诱导宿主免疫反应,并且可以在无辅助病毒系统中产生。AAV 载体可以在长期培养中有效地转导 CD34+ 细胞(查特杰(Chatterjee)等人,1999)。

[0166] 整合载体,例如逆转录病毒或慢病毒,通常被用于基因疗法中,然而,这些载体的随机整合,以及一些诱导基因的致癌性质,造成了癌症形成的风险。为此,优选非整合方法,例如腺病毒、杆状病毒或者能够附加型表达的质粒的瞬时转染。优选以非整合方式将基因引入细胞基因组中的载体。允许引入的基因瞬时表达的病毒载体也是优选的。在宿主细胞中具有短生命周期的载体也是优选的。

[0167] 天然存在的组成型启动子控制必要细胞功能的表达。因此,处于组成型启动子控制下的基因在所有细胞生长条件下表达。示范性组成型启动子包括用于以下编码某些组成型或“看家(housekeeping)”功能的基因的启动子:次黄嘌呤磷酸核糖基转移酶(hypoxanthine phosphoribosyl transferase, UPRT)、二氢叶酸还原酶(dihydrofolate reductase, DHFR)(斯查夫曼(Scharfmann)等人,1991,美国国家科学院院刊(Proc. Natl. Acad. Sci. USA), 88:4626-4630)、腺苷脱氨酶、磷酸甘油激酶(phospho glycerol kinase, PGK)、丙酮酸激酶、磷酸甘油变位酶、肌动蛋白启动子(赖(Lai)等人,1989,美国国家科学院院刊, 86:10006-10010),以及本领域技术人员已知的其它组成型启动子。此外,许多病毒启动子在干细胞中组成性地起作用。这些启动子包括:SV40 的早期和晚期启动子;莫洛尼白血病病毒和其它逆转录病毒的长末端重复(long terminal repeats, LTRs);以及单纯疱疹病毒的胸苷激酶启动子等等。因此,上文提到的任何组成型启动子都可以用于控制异源基因插入物的转录。

[0168] 处于诱导型启动子控制下的基因仅在诱导剂存在下表达或在诱导剂存在下较高程度表达(例如,处于金属硫蛋白启动子控制下的转录在某些金属离子存在下大幅增加)。

诱导型启动子包括应答元件(responsive element, RE),它在结合于其诱导因子时刺激转录。举例来说,存在血清因子、类固醇激素、视黄酸和环 AMP 的 RE。含有特定 RE 的启动子可以经过选择以便获得诱导型应答,并且在一些情形中,RE 本身可以连接到不同的启动子,由此赋予重组基因可诱导性。因此,通过选择适宜的启动子(组成型对诱导型,强对弱),有可能控制一种试剂在经过遗传修饰的细胞中的存在和表达水平。考虑到上文公开的因素,这些用于基因插入物表达的因子的选择和优化被认为在本领域普通技术人员的技能范围内,无需过多实验。

[0169] 产生病毒载体和使用病毒载体进行基因疗法的方法是本领域中众所周知的。举例来说,参看,基因转移:DNA 和 RNA 的递送和表达实验手册(Gene transfer:delivery and expression of DNA and RNA:a laboratory manual)(CSHL 出版社,2007);用于基因疗法的病毒载体:方法和实验指南(Viral vectors for gene therapy:methods and protocols)(胡马纳出版社,2003);基因疗法实验指南:第 2 卷:基因转移载体的设计和表征(Gene Therapy Protocols:Volume2:Design and Characterization of Gene Transfer Vectors)(胡马纳出版社,2008);以及基因和细胞疗法:治疗机制和策略(Gene and cell therapy:therapeutic mechanisms and strategies)(CRC 出版社,2008)。这些文献以引用的方式并入本文中。

[0170] 蛋白质转导

[0171] 蛋白质转导可以用作病毒载体的替代选择用于将蛋白质递送到干细胞中。蛋白质转导是蛋白质从外部环境内化进入细胞中。这一过程依赖于少数蛋白质和肽能够穿透细胞膜的固有特性。这些分子的转导特性可以赋予与其一起以融合物形式表达的蛋白质。根据本发明可以用于多肽的蛋白质转导的肽的实例包括以下:

[0172] 触足肽:触足基序是来源于果蝇同源异型蛋白(一类参与发育过程的转活化因子)家族。这些蛋白质通过排列成三螺旋序列的 60 个氨基酸的羧基末端区域(称为同源异型结构域)识别并结合 DNA。触足(AntpHD)的同源异型结构域能够转位跨过神经元膜并被运送到核。

[0173] 单纯疱疹病毒 VP22 蛋白:1 型单纯疱疹病毒(HSV-1)VP22 蛋白是一种结构多肽,它形成位于成熟病毒粒子的包膜与衣壳区之间的病毒被膜的主要组分。它是一种小的碱性蛋白,约 38kDa 大,由 UL49 基因编码。

[0174] HIVTAT 蛋白转导结构域:经显示,HIV-1 转活化基因产物 TAT 是潜伏期 HIV 中的转录调控因子,并且对于 HIV 复制极为重要。它是一种由分别具有 72 个和 14 个氨基酸的两个外显子构成的具有 86 个氨基酸的蛋白质。

[0175] Chariot 蛋白:Chariot 是一种 2843 道尔顿的肽并且与关注的蛋白质形成共价复合物(艾莫夫有限公司(Active Motif, inc.))。

[0176] 可以用于蛋白质转导的肽的其它实例包括富精氨酸肽(15)和 +36GFP (16)。

[0177] 制备用于蛋白质转导的肽的方法是本领域中众所周知的。举例来说,参看,细胞穿透肽手册(Cell-penetrating peptides:Handbook of)(CRC 出版社,2006)、细胞渗透肽:方法和应用(Cell-penetrating peptides:processes and applications)(CRC 出版社,2002)、蛋白质转导:递送重组 Tat-PTD 融合蛋白到目标细胞中(Protein transduction:delivery of recombinant Tat-PTD fusion proteins into target

cells)(特罗瑟姆大学(University of Tromsø),医学生物学研究所(Institute of Medical Biology),2000)及蛋白质分析和纯化:实验技术(Protein analysis and purification:benchtop techniques)(施普林格出版公司(Springer),2005)。这些文献以引用的方式并入本文中。

[0178] 通过参照以下实验详情将更好地理解本发明,但本领域技术人员将易于了解,详细说明的具体实验只是对以上权利要求书中较完整描述的本发明的说明。

[0179] 方法

[0180] 人 cd34+ 细胞的分离和 SALL4- 慢病毒感染

[0181] 购买正常人骨髓 CD34+ 细胞(奥赛尔公司(AllCells),美国加利福尼亚州爱莫利维尔市(Emeryville, CA, USA)),并且在 StemSpan SFEM(干细胞技术公司(Stemcell Technologies),加拿大卑诗省温哥华市(Vancouver, BC, Canada))中产生单细胞悬浮液,保持 24 小时。接下来,将 10^5 个细胞/孔接种到 12 孔板中。在含有 10%FBS 和 1%青霉素/链霉素(pen/strep)(吉博公司(Gibco),美国加利福尼亚州卡尔斯巴德(Carlsbad, CA, USA))的 StemSpan SFEM 中培养细胞。另外,培养基补充有 100ng/ml FLT-3、100ng/ml TPO 和 100ng/ml SCF(普鲁派克公司(ProSpec),以色列雷霍沃特(Rehovet, Israel))。在 37℃ 下,将 MOI 介于 10 到 20 之间的 SALL4 慢病毒粒子添加到骨髓细胞中。对于对照物,将 MOI 类似的仅含 GFP 的慢病毒粒子添加到骨髓细胞中。感染细胞 2.5 小时,并在培养基中恢复 24 小时。次日,再次感染细胞 2.5 小时,漂洗并接种用于实验扩增。

[0182] SALL4 转导的人 CD34+ 细胞的扩增

[0183] 在标准培养条件(含有 10%FBS 和 1%青霉素/链霉素并补充有 100ng/ml FLT-3、100ng/ml TPO 和 100ng/ml SCF 的 StemSpan SFEM)下,在 12 孔板中扩增用 SALL4 转导的 CD34+ 细胞。此外,在更严格的条件下培养细胞,在所述条件下,降低重组合细胞因子的浓度以观察用 SALL4 转导的细胞是否仍能够存活和扩增。在一个实验中,在含有不到 50%细胞因子(50ng/ml FLT-3、50ng/ml TPO 和 50ng/ml SCF)的培养基中培养细胞。在另一试验中,在含有不到 75%生长因子(25ng/ml FLT-3、25ng/ml TPO 和 25ng/ml SCF)的培养基中培养细胞。监测细胞两个月,并利用明视场和荧光显微术进行观察。

[0184] 用 SALL4 转导的从患者样品分离的 CD34+ 细胞的扩增

[0185] 从位于美国纽约州石溪(Stony Brook, NY, USA)的血液骨髓干细胞实验室(Blood Marrow Stem Cell Laboratory)获得外周血干细胞。使用 CD34 人微珠试剂盒(Microbead kit)和 MiniMACS 分离柱(美天旎生物技术有限公司(Miltenyi Biotec Inc.),美国加利福尼亚州奥本(Auburn, CA, USA))从干细胞池中分离 CD34+ 细胞。在分离出 CD34+ 细胞后,用慢病毒转染细胞并以与先前所提到相同的方式进行培养。

[0186] 骨髓细胞的 CFU 分析

[0187] 在 4° 的冰箱中解冻含有 MethoCult®(干细胞技术公司;加拿大卑诗省温哥华市)培养基的管。次日早晨,对这些管进行涡旋以确保所有组分都充分混合。随后,按 10 倍所需最终浓度制备用 SALL4 转导或用 GFP 转导的 CD34+ 细胞。制备每毫升 1×10^6 个细胞的细胞悬浮液,并在不同浓度(2-4 倍)下,每种浓度配制两份。将 0.3mL 细胞添加到 3mL 的 MethoCult®培养基(干细胞技术公司;加拿大卑诗省温哥华市)进行一式两份培养。再对管进行涡旋以确保所有细胞和组分都充分混合,并使其静置 5 分钟以使气泡消散。使用连

接到 3cc 注射器的 16 号钝头针将细胞和 **MethoCult**® 培养基分配到培养皿中。每一个 35mm 的培养皿分配 1.1mL 细胞。通过小心地倾斜并旋转每一培养皿来使甲基纤维素培养基和细胞均匀地分布。将两个培养皿放入 100mm 的皮氏培养皿 (petri dish) 中, 并且还增加第三个含有 3mL 无菌水的未覆盖的 35mm 培养皿。随后, 将全部 3 个培养皿覆盖在 100mm 的皮氏培养皿内。在 37℃ 以及 5%CO₂ 和 ≥ 95% 湿度下, 培育细胞 14 到 16 天。利用明视场和荧光显微术观察 BFU-E、CFU-GM 和 CFU-GEMM 集落。此外, 在将细胞接种于 **MethoCult**® 培养基中后 10 到 18 天, 在显微镜下对 CFU 计数。具有超过 100 个细胞的集落计为阳性集落。

[0188] 细胞的流式细胞术和表型分析

[0189] 用偶联藻红蛋白 (Phycoerythrin, PE) 的抗 CD34 抗体进行 FLOW, 使用偶联别藻蓝蛋白 (allophycocyanin, APC) 的抗 CD38 抗体和偶联 PerCP-Cy5.5 的抗 CD45 抗体进行细胞分选 (BD 生物科技公司 (BDBiosciences), 美国新泽西州富兰克林湖 (Franklin Lakes, NJ, USA))。使用 CD45-PE 抗体确定在 NOD-SCID 小鼠骨髓中人类细胞的存在。利用 CD15-APC 或 CD33-APC 抗体追踪髓样细胞, 并且利用 CD19-PE 或 CD3-PE 抗体追踪淋巴样细胞。

[0190] 细胞计数和生长曲线实验

[0191] 使用血细胞计数器, 每 24 小时对用 SALL4A 转导的细胞、用 SALL4B 转导的细胞和对照细胞进行计数, 持续 7 天。在计数前, 用 200mL 的移液管尖小心地抽吸细胞以便将细胞簇解离成个别细胞。

[0192] 长期培养分析

[0193] 在整体或限制性稀释条件下, 在 **MethoCult**® 培养基中进行使用人 CD34+ 细胞 LTC-IC 分析。为了计算 LTC-IC 的总数, 由二次培养测定 LTC-IC 的频率。随后, 将这一计算的数量乘以在 4 周的原代长期培养后存在的细胞总数。

[0194] NOD/SCID 小鼠再群体化细胞 (SRC) 分析

[0195] 第 0 天, 将 4×10^3 个、 8×10^3 个、 2×10^4 个或 4×10^4 个 2 天大的仅表达 GFP 的人 CD34+ 骨髓细胞或 14 天大的经培养的表表达 SALL4A 或 SALL4B 的人 CD34+ 骨髓细胞, 以及 2×10^5 个 CD34- 辅助细胞注射到经过照射 (2, 5Gy) NOD-LtSz-scid/scid (NOD-SCID) 小鼠中。移植后 7 周, 通过流式细胞术分析外周血样品中 CD45+ 细胞的存在。当在小鼠外周血细胞中检测到至少 0.5% 的 CD45+ 人类细胞时, 将小鼠评为对人类定植呈阳性。由产生 37% 的阴性小鼠的测试细胞浓度的倒数测定干细胞起始的细胞频率。动物实验是根据由石溪大学 (Stony Brook University) 的机构动物管理与使用委员会 (Institutional Animal Care and Use Committee, IACUC) 批准的研究人员方案执行。

[0196] 连续移植研究

[0197] 在移植后 16 周, 从高度定植的原代接受者小鼠的胫骨和股骨采集小鼠 BM 细胞。在利用溶解缓冲液去除红细胞后, 将来自每一接受者小鼠的半数 BM 细胞移植到一只经过亚致死照射 (2.5Gy) 的二次 NOD/SCID 小鼠中。移植后 5 周, 如所述, 通过流式细胞术分析在二次接受者小鼠的外周血中人 CD45+ 细胞的百分含量。以相同方式进行第三次移植并在移植后 10 周, 对骨髓细胞进行流式细胞术以分析 CD45+ 细胞、CD33+ 细胞、CD19+ 细胞和 CD3+ 细胞。

[0198] His-TAT-SALL4A/B 的表达和纯化

[0199] 对于在大肠杆菌中表达 TAT-SALL4, 用限制性内切核酸酶 SalI 和 NotI 双重消化人 SALL4B 基因, 并连接到 pTAT-pET28b 载体中, 所述载体是带有 T7 启动子、卡那霉素 (kanamycin) 抗性和 pTAT 肽 / 蛋白质转导结构域 (Protein Transducing Domain, PTD) 的 6×His-taq 蛋白质表达载体。随后, 将具有正确基因序列的质粒转化到大肠杆菌菌株 BL21 (DE3)。在 0.1mMIPTG 诱导作用下, 表达融合 6×His 的人 SALL4B 达 3 小时, 随后通过使用 Ni-NTA 亲和柱 (凯杰公司 (Qiagen), 美国加利福尼亚州巴伦西亚 (Valencia, CA, USA)) 进行纯化。基于 SDS-PAGE 和 LC-MS/MS 进一步测定纯化的 His-TAT-SALL4B 的纯度。

[0200] SDS-PAGE 和蛋白质免疫印迹

[0201] 将蛋白质装载到 12.5% 的一维 SDS-PAGE 上用于蛋白质分离, 随后用考马斯亮蓝 R-250 进行染色, 并在 10% 甲醇 / 7% 乙酸中脱色。对于蛋白质免疫印迹分析, 将 SDS-PAGE 转移到聚 (偏二氟乙烯) (poly-(vinylidene difluoride), PVDF) 膜 (密理博公司 (Millipore), 美国马萨诸塞州比尔里卡 (Billerica, MA, USA))。转移后, 在 4℃ 下用含 5%w/v 脱脂奶粉的 TBS/0.1%Tween20 使膜饱和过夜, 随后与一次抗体一起在 4℃ 下培育过夜。抗人 SALL4 的一次抗体和 6×His 标签是购自艾宾公司 (Abeam) (英国剑桥 (Cambridge, UK))。用 TBS/0.1%Tween20 洗涤三次后, 将膜与偶联过氧化物酶的二次抗体 (艾宾公司; 英国剑桥) 的溶液一起培育。在室温下培育 1 小时后, 用 TBS/0.1%Tween20 洗涤膜三次, 并通过使用 ECL 底物 (密理博公司) 展开膜印迹。

[0202] 胶内消化和 LC-MS/MS

[0203] 切割凝胶带并用胰蛋白酶进行胶内消化。将肽再悬浮于 50 μl 的 0.1% 甲酸 / 2% 乙腈中。将 10 μl 肽注射到 Orbitrap 中, 在石溪大学的蛋白质组学中心 (Proteomics Center) 进行 1DLC/MS/MS 分析。

[0204] 昆虫细胞的生长和维持

[0205] 使 Sf9 昆虫细胞在补充有 10%FBS (吉博公司) 和 1% 抗生素 / 抗霉菌素 (Cellgro) 的 Sf900-II-SFM (吉博公司) 培养基中生长。使用前, 使培养基滤过 ExpressPLUS0.22 μm 过滤器 (密理博公司)。27℃ 下, 在旋转瓶 (惠顿公司 (Wheaton)) 或组织培养摇瓶 (费舍尔公司 (Fisher), 美国宾夕法尼亚州匹兹堡 (Pittsburgh, PA, USA)) 中将细胞数量维持在每毫升介于 0.5 与 5.0×10^6 个细胞之间,

[0206] 用含有 His-TAT SALL4B 构建体的杆状病毒感染 Sf9 细胞

[0207] 用含有 His-TAT SALL4B cDNA (已经在 Sf9 细胞中新鲜扩增) 的杆状病毒感染每毫升计数达到 2.5×10^6 个细胞的细胞。在惠顿公司的 250mL 旋转瓶中, 最终体积为 100mL。感染的细胞在 27℃ 下继续培育 4 天, 随后采集。

[0208] Sf9 细胞的溶解以及 His-TAT hSALL4B 的纯化

[0209] 以 500×g 离心细胞 10 分钟。保存上清液作为病毒储备液。将细胞团粒再悬浮于 20mL 含有 50mM NaH₂PO₄ (pH8.0)、0.3M NaCl、蛋白酶抑制剂混合液、PMSF 和 1%NP40 (IGEPAL) 的溶解缓冲液中。在冰上培育 30 分钟后, 通过在 4℃ 下以 10,000g 离心 15 分钟来使溶解产物变澄清。

[0210] 将澄清的溶解产物添加到用溶解缓冲液洗涤了 4 次的 Ni-NTA 珠粒 (凯杰公司) 中。每 4mL 溶解产物添加 200 μL 珠粒。为了限制非特异性结合, 添加 10mM 咪唑。4℃ 下, 在旋转器上培育溶液达 2 小时, 随后将其装载到重力流式柱上并收集流出物。首先用 10mL 含 20mM

咪唑的溶解缓冲液洗涤珠粒,随后用 10mL 含 80mM 咪唑的溶解缓冲液进行洗涤。用溶解缓冲液和 250mM 咪唑执行洗脱,以 1.5mL 洗脱份收集,直到 10mL。通过 8%PAGE-SDS 以及使用抗 SALL4 (亚诺法公司(Abnova))的 1:2000 稀释液进行的免疫印迹来分析所有溶解产物和洗脱份。His-TAT SALL4B 通常在洗脱份 2 和 3 中洗脱。汇集这些洗脱份并在 Slide-A-Lyzer 盒(皮尔斯公司(Pierce);美国伊利诺伊州罗克福德(Rockford, IL, USA))针对 PBS 透析过夜。

[0211] 统计学分析

[0212] 结果报导为平均值 $\pm$ 标准差。 $p < 0.05$ 的值被认为是统计学上显著的。

[0213] 实验详情

[0214] 实例 1

[0215] 过表达 SALL4 的人 HSC 的离体扩增

[0216] 为了解用 SALL4 诱导的 HSC 扩增的幅度,对在慢病毒转导后人骨髓 CD34 阳性细胞中 SALL4 过表达的影响进行研究。将 SALL4A 和 SALL4B cDNA 并入带有 GFP 报告基因的载体中,所述报告基因有助于转导细胞的分离和追踪。SALL4A 和 B 是由内部剪接产生的两种剪接同功异型物(图 1)。

[0217] 用含有 h-FLT-3、h-SCF 和 h-TP0 的生长培养基培养正常人 CD34+ 细胞 2 天。48 小时后,将细胞分为具有 3×10^4 个细胞的六个组,并放入 12 孔板的个别孔中。随后,用 SALL4A、SALL4B 或 GFP (对照物) 人慢病毒转导细胞,保持 2.5 小时。2 小时后,使细胞在生长培养基中恢复。24 小时后,再次用上文提到的慢病毒转导细胞,保持 2.5 小时。次日,观察这些细胞中的 GFP 阳性细胞。注意到,约 40% 的细胞为发荧光的(图 2)。

[0218] 第二次慢病毒感染后,在接下来的 14 天里每天观察细胞。用明视场和荧光显微术定性地观察细胞。在前 48 小时期间,用 SALL4A 或 SALL4B 转导的细胞扩增 10 倍,并且有规则地形成 GFP 阳性细胞簇。相比之下,仅用 GFP 转导的细胞扩增约 2 倍,但没有形成许多簇。而是观察到单个细胞或很少的小 GFP 阳性簇。7 天后,扩增细胞看起来只在用 SALL4 转导的细胞中观察到(图 3A)。8 到 11 天后,在完全生长培养基中,用 SALL4A 或 SALL4B 转导的细胞扩增 200 到 400 倍(图 4 和 5)。另一方面,仅用 GFP 转导的细胞最多扩增 5 倍(图 4 和 5)。此外,在慢病毒感染后 16 天,用 SALL4 转导的细胞继续存活并扩增。这些细胞在培养的前 2 到 3 周里能够扩增 800 到 1000 倍,同时没有显著成熟,而大多数对照细胞经历了死亡或分化(图 6)。因此,仅在用 SALL4 转导的细胞中观察到细胞扩增。重复扩增实验三次。在所有试验中,用 SALL4A 或用 SALL4B 转导的人 CD34+ 细胞能够在液体培养基中生长,保持其未分化的外观,并且继续生长超过 2 个月。利用从来自患者样品的外周血干细胞分离的 CD34+ 细胞进行类似实验。最初的发现证实,这些细胞的成功存活和扩增也是可能的。

[0219] 利用从动员的外周血干细胞分离的 CD34+ 细胞进行类似实验,这些外周血干细胞是从患者收集并经历低温保存。来自不同患者的三个样品在 CD34+ 细胞扩增方面显示出期望的结果(图 3B)。用 SALL4A 或 SALL4B 转导的细胞扩增约 130 倍,而对照细胞最多扩增 12 倍(图 7A)。此外,当这些用 SALL4 转导的细胞生长 14 天时,CD34+/CD38- 群对于 SALL4A 增加 369 倍并且对于 SALL4B 增加 384 倍(图 7B),并且维持干细胞免疫表型($n=4$) (图 8)。细胞培养 31 天时,用 SALL4 转导的扩增细胞仍保持祖细胞或干细胞的免疫表型($n=2$) (图 7C) 以及原始细胞形态(图 7D)。在扩增结束时,约 90% 的细胞表达 GFP。相比之下,在感染

后两周以后,对照细胞停止增殖,或不再存活。用 SALL4 诱导的细胞继续扩增,并且 CD34+/CD38- 或 CD34+/CD38+ 比率类似于开始培养时初始细胞的比率(图 7)。

[0220] 使用多种 CD34+ 细胞来源重复这些扩增实验至少 12 次(表 1)。在所有试验中,用 SALL4A 或 SALL4B 转导的人 CD34 细胞能够在液体培养基中扩增超过 2 个月,并且呈现与 31 天大的细胞类似的细胞群百分含量(15%CD34+/CD38-、17%CD34+/CD38+),如通过流式细胞术所证实。

[0221] 表 1. CD34+ 细胞的来源

[0222]

CD34+细胞的来源	来源
奥赛尔斯公司	27 岁; 女性; 菲律宾裔/白种人
奥赛尔斯公司	27 岁; 女性; 菲律宾裔/白种人
奥赛尔斯公司	32 岁; 男性; 白种人
奥赛尔斯公司	32 岁; 男性; 白种人
奥赛尔斯公司	32 岁; 男性; 白种人
SBUMC 骨髓移植实验室 (Bone Marrow Transplant Lab)	保密
SBUMC 骨髓移植实验室	保密
SSUMC 骨髓移植实验室	保密
奥赛尔斯公司	22 岁; 男性; 白种人
奥赛尔斯公司	22 岁; 男性; 白种人
奥赛尔斯公司	22 岁; 男性; 白种人
奥赛尔斯公司	22 岁; 男性; 白种人

[0223] 实例 2

[0224] 单个用 SALL4 诱导的 CD34+ 细胞在培养中容易地扩增

[0225] 从第 5 到 14 天,可以在用 SALL4A 或 SALL4B 转导的细胞培养物中容易地观察到较大的 GFP 阳性细胞簇。为了解这些簇是否能用作种子进行进一步扩增,从亲本板拣取单个细胞簇,并放入 48 孔板的个别孔中。通过用 100 μ l 移液管尖抽吸,小心地解离细胞簇。次日(第 6 到 15 天),新的小 GFP 阳性细胞簇又开始形成(图 9)。用完全生长培养基使这些簇扩增 5 天,并且它们开始以与在初始亲本板中所见类似的速率生长(图 9)。因此,有可能选择出用 SALL4 转导的细胞簇并使其在 7 天内扩增 200 到 400 倍。

[0226] 为了进一步说明变老的用 SALL4 转导的 HSC 集落的存活和扩增能力,将 18 天大的细胞放入细胞离心机中并进行吉姆萨染色以观察其形态。有趣的是,超过 90% 的细胞展现类母细胞的形态,具有大核和稀少的胞浆(图 10)。这些细胞类似于未分化的细胞。此外,变老的用 SALL4 转导的 HSC 在 Methocult 中培养并针对 CFU 评分。用 SALL4 转导的 GFP 阳性 HSC 能够分化成不同 CFU 集落,包括 BFU-E、CFU-GM 和 CFU-GEMM(图 11 和 12)。因此,用 SALL4 诱导的 HSC 细胞仍能够形成多种血液谱系,与野生型 CD34 细胞类似。

[0227] 实例 3

[0228] 人造血干细胞在严格条件下离体用 SALL4 诱导的扩增

[0229] 为了证实用 SALL4 诱导的 HSC 相比对照物能够以较快的速率和较高的体积扩增,

进行细胞培养实验,在这一实验中,显著地降低重祖细胞因子浓度。如果用 SALL4 转导的 HSC 能够在这些更为严格的条件下成功地生长,那么可以巩固我们的发现:SALL4 与 HSC 维持未分化的增殖状态和阻止细胞分化有关。

[0230] 进行细胞培养实验,在这些实验中,对于 SALL4 诱导的 HSC 和对照细胞,使重组生长因子浓度降低 50% 或 25%。在第一周期间,监测细胞生长。有趣的是,在超过 6 天的培养中,用 SALL4A 或 SALL4B 转导的细胞仍能够存活并扩增(图 13 和 14)。相比之下,到第 6 天时,对照细胞经历了细胞死亡,并且不能观察到存活的集落。此外,在这些条件下,用 SALL4 转导的人 CD34⁺ 细胞能够在液体培养基中维持其未分化的外观,并且继续生长超过 2 个月。

[0231] 这提供了第一个证据证明,在降低必要重祖细胞因子浓度的严格生长条件下,用 SALL4 诱导的 HSC 仍具有在 100% 浓度下培养的细胞的存活和扩增能力。此证实了 SALL4 对于将 HSC 维持在未分化的增殖状态起到关键作用。

[0232] 此首个数据说明,用 SALL4 转导骨髓 CD34⁺ 细胞提供了一种成功地扩增 HSC 细胞的方式。不存在 SALL4 过表达的 HSC 的数量因分化或死亡而减少,导致净 HSC 减少。相比之下,对于过表达 SALL4 的 HSC,许多克隆能够在培养中存活并扩增。在整个培养中,展现净 HSC 扩增,具有许多正在扩增的细胞簇(图 15)。这一数据增长了有关 HSC 如何自更新和扩增的知识。此外,这一知识可以转换成在临床上扩增 HSC 的有用数量的实验方案,用于针对血液系统病症的骨髓移植和靶向基因疗法。

[0233] 在其它实验中,用 SALL4 转导的细胞在具有生长因子 SCF、TPO 和 FLT-3L 的各种组合的培养基中生长。培养 14 天后,注意到,用 SALL4 转导的细胞的生长和存活与 FLT-3L 无关,部分地取决于 TPO,并且取决于 SCF (表 2)。

[0234] 表 2. 用 SALL4 转导的细胞对某些生长因子的依赖性

[0235]

细胞因子	SALL4A	SALL4B
FLT-3	无关	无关
TPO	部分相关	部分相关
SCF	相关	相关

[0236] 实例 4

[0237] 变老的用 SALL4 诱导的 HSC 的形态和表型分析

[0238] 为了证实用 SALL4 诱导的 HSC 在数周的老化后仍具有祖细胞特征,在不同时间点时进行细胞的细胞离心和吉姆萨染色。有趣的是,在慢病毒感染后 16 天和 31 天,细胞形态高度类似于祖细胞,具有大核和稀少的胞浆(图 10)。此外,还在两个不同时间点时进行 FLOW,并显示在对照细胞中与在变老的用 SALL4 诱导的细胞中 CD34⁺/CD38⁻ 细胞的比率类似(图 16)。这是一个关键的发现,因为这些细胞甚至在培养数周后仍以较高百分含量呈现未分化的祖细胞状态。

[0239] 先前的研究已经说明,SALL4 在细胞存活和细胞凋亡中起到重要作用。此外,SALL4 能够结合于在 HSC 扩增和自更新中起关键作用的众多基因(包括 BMI1、WNT1 和 TGFB1)的启动子区(图 17)。

[0240] 实例 5**[0241] SALL4 的过表达将抑制在 32D 细胞系中的粒细胞分化**

[0242] 为进一步确定 SALL4 如何影响干细胞生长,将研究扩充到髓样祖细胞系 32D,在这一细胞系中,其正常表达。32D 细胞当维持在 IL-3 中时以未分化的母细胞形式增殖,但在用 G-CSF 刺激时分化成为成熟的嗜中性粒细胞。

[0243] 用含有 IL-3 的生长培养基培养 32D 细胞 3 天。第三天,将细胞分为具有 5×10^4 个细胞的六个组,并放入 12 孔板的个别孔中。随后,用 SALL4A、SALL4B 或 GFP (对照物) 人慢病毒转导细胞,保持 2 小时。2 小时后,使细胞在生长培养基中恢复。24 小时后,再次用上文提到的慢病毒转导细胞。次日,观察这些细胞中的 GFP 阳性细胞。注意到,约 30% 的细胞为发荧光的。使细胞在补充有 IL-3 的生长培养基中扩增 48 小时。48 小时后,收集细胞,洗涤并再接种于含有生长培养基的新孔中,所述生长培养基不含 IL-3,但补充有 G-CSF。

[0244] 在接下来的七天里,每天用明视场和荧光显微术观察细胞,同时用血细胞计数器计数。在第 1 到 7 天中,即使从生长培养基中去除了 IL-3,用 SALL4A 或 SALL4B 转导的细胞仍继续以稳态扩增。在 G-CSF 中培养 3 天后,32D-SALL4A 或 32D-SALL4B 的增殖速率是对照对应物的 3 倍或 6 倍(图 18)。当去除 IL-3 并用 G-CSF 代替时,对照 32D 细胞在 G-CSF 中 5 到 6 天后死亡,但 32D-SALL4A 和 32DSALL4B 细胞在培养中无限地生长。此外,这些细胞的表现与未经过修饰的 32D 细胞相同(图 18 和 19)。SALL4A 或 SALL4B 的表达允许细胞以未分化的状态继续生长。对照 32D 细胞展现适宜的粒细胞成熟,而 32D-SALLA 或 32D-SALL4B 细胞不显示显著的粒细胞成熟,并且其形态与 32D-SALL4 亲代群极为相似。32D 细胞当维持在 IL-3 中时以未分化的母细胞形式增殖(图 20iv),但在用 G-CSF 刺激时分化成为成熟的嗜中性粒细胞(图 20vi)。用 SALL4 转导的 32D 细胞在不存在 IL-3 时无限地生长,并且在给予 G-CSF 时保持未分化的母细胞形态(图 20v)。这一研究指出,可以通过 SALL4 的组成型表达来阻止造血干细胞的分化。

[0245] 实例 6**[0246] SALL4A 和 SALL4B 的体外相互作用**

[0247] 为了确定 SALL4A 或 SALL4B 蛋白在体外是形成同源二聚体还是异源二聚体,用 HA 或 His (6 个组氨酸残基)对 SALL4A 或 SALL4B 进行标记,并执行抗 HA 下拉分析。如图 21A 和 B 中所示,表达 HA-SALL4A 和 His-SALL4B 或 HA-SALL4B 和 His-SALL4A 蛋白。执行抗 HA 下拉分析。如图 21C 中可见,His-SALL4A 或 His-SALL4B 被 HA-SALL4A 或 HA-SALL4B 下拉。

[0248] 实例 7**[0249] 用 SALL4 诱导的 CD34⁺ 细胞的体外功能分析**

[0250] 为了进一步研究用 SALL4 转导的 HSC 在无过量特定细胞因子存在下的增殖潜能,进行集落形成单位(colony-forming unit,CFU)分析。用于此研究的 CFU 祖细胞最初为 GFP 阳性的,表明 SALL4 蛋白的表达(图 11),并且在含有细胞因子 SCF、TPO 和 FLT-3L 的培养基中培养至少 1 个月。在不含这些细胞因子的 methocult 培养基中进行的 CFU 分析揭露,这些细胞可以形成各种集落,包括 CFU-GM、CFU-GEMM 和 BFU-E (图 22A)。此外,当在起始 CFU 分析后 18 天进行计数时,1 个月大的经过转导的细胞能够形成与 2 天大的用 GFP 转导的对照细胞类似数量的 CFU 集落(图 22B 和 22C)。不希望受任何科学理论的束缚,此研究可以指出,SALL4 与其它细胞因子结合作用来阻止 HSC 分化,并且可能需要其它细胞因子来抑制分

化。

[0251] 为了进一步测试 SALL4 的转移是否会引起人造血祖细胞的长期扩增,将用 SALL4A 和用 SALL4B 转导的细胞培养 1 个月。培养 31 天后,与对照物相比较,总 CD34⁺ 细胞数量对于 SALL4A 来说存在 1780 倍增加,并且对于 SALL4B 来说存在 1463 倍增加(图 23A)(对照细胞在 10 到 12 天后停止扩增)。此外,在 1 个月,对于 LTC-IC 的总数,相比对照物,用 SALL4 转导的细胞对于 SALL4A 来说显示出 9.32 倍增加并且对于 SALL4B 来说显示出 8.88 倍增加(图 23B)。总的说来,用 SALL4A 转导的细胞相对于对照物具有 16700 倍的总 CD34⁺/CD38⁻ 干细胞扩增倍数,而用 SALL4B 转导的细胞显示出 13300 倍增加,不过两种 SALL4 同功异型物之间不存在统计学显著的差异(图 23C)。

[0252] 实例 8

[0253] 扩增伴随体内干细胞再群体化的增强

[0254] 将异种移植分析变成测试受 SALL4 介导的细胞是否能够压倒骨髓小生境中的调控机器来控制其分化、再群体化能力和干细胞输出。表型分析揭露,感染后 4 周,用 SALL4A (n=9)和用 SALL4B 转导(n=10)的细胞能够将阳性细胞定植到 NOD/SCID 小鼠体内(图 23D)。此外,实验动物(n=8)在注射后 4 周呈现髓样(CD15⁺)和淋巴样(CD19⁺)谱系(图 23E),指示了髓样/淋巴样分化过程。使用流式细胞术分析来自注射后 15 周的动物(n=4)的骨髓,并且仍展现三个谱系分化,包括 CD3⁺/CD19⁺ 淋巴样谱系、CD15⁺ 髓样谱系和血型糖蛋白-4A⁺ 红系谱系(图 24)。

[0255] 为了进一步确定用 SALL4 诱导的细胞是否具有长期可定植特性,用从注射有人 CD34⁺ 细胞的原代动物采集的骨髓进行二次和三次移植。Flwo 分析证实,注射后 8 周,动物骨髓仍具有 CD45⁺ 细胞(2.74%),但未得到大量 CD34⁺ 细胞(<0.24%),从而显示无致白血病作用。

[0256] 此外,CD45⁺ 群对 CD3⁺ (2.97%)/CD19⁺ (0.89%) 淋巴样谱系、CD15⁺ (14.1%) 髓样谱系和血型糖蛋白-4A⁺ 谱系(5.6%)呈阳性。这显示,将骨髓从一只动物成功地移植到另一只动物是可能的,并且细胞可以在骨髓小生境中适当地分化。此外,用从二次移植动物采集的骨髓进行第三次移植。移植后 10 周,动物仍在其骨髓内展现 CD45⁺ (3.29%) 细胞(图 23F 和 23G)。有关 CD45⁺ 群的进一步分析证实,细胞对髓样谱系和淋巴样谱系呈阳性,与第二次移植的细胞类似(图 23F 和 23G)。

[0257] 培养 4 周后,SALL4A 或 SALL4B 增加人类细胞再群体化成 NOD/SCID 的潜能的能力由移植有用 SALL4A 或 SALL4A 转导的细胞的小鼠中嵌合体水平的大幅增大得到证实。带有 SALL4A 的 CD45⁺ 细胞的比例为 3.84%(0.34-12.89%;n=12)并且带有 SALL4B 的 CD45⁺ 细胞的比例为 2.98% (0.41-9.41%;n=12),与对照细胞的 0.52% (0.03-1.18%;n=12) 形成对比(图 23H)。此外,通过检测在 14 周(n=6)或 19 周(n=3)时小鼠外周血中的人 CD45⁺ 细胞,也显而易见用 SALL4A 或 B 转导的细胞的长期定植。

[0258] 为了定量地测量用 SALL4 转导的 CD34⁺ 细胞的作用,进行限制性稀释实验以测定 NOD/SCID 小鼠再群体化细胞(SRC)的频率。注射后 7 周, SRC 频率由 1/19,200 个 CD34⁺ 细胞(由 +/-s.e. 界定的范围:14,100-25,900)增加到分别对于 SALL4A 为 1/7,100 (由 +/-s.e. 界定的范围:5,200-9,800)和对于 SALL4B 为 1/9,100 (由 +/-s.e. 界定的范围:6,700-12,400)。通过考虑 SRC 频率的增加和总细胞数量,与未培养的细胞相比较,总 SRC

含量对于 SALL4A 来说扩增 1080 倍并且对于 SALL4B 来说扩增 844 倍($n=72$) (图 23I)。不希望受任何科学理论的束缚,这些结果可以表明,仅在 HSC 中过表达 SALL4 不能压倒参与体内干细胞输出控制的调控机器。

[0259] 实例 9

[0260] TAT-SALL4B 蛋白诱导 CD34⁺ 细胞扩增

[0261] SALL4 的慢病毒表达是极为有效的,但其临床应用并不理想,因为难以控制转基因在体内表达的水平 and 持续时间以及插入性白血病形成(insertional leukemogenesis)的可能。为了使用不同的方法来证实 SALL4 在 HSC 扩增中的作用,在大肠杆菌中表达 TAT-6 $\times$ His-SALL4B 蛋白,并使用 Ni-NTA 琼脂糖进行纯化(图 25A 和 25B)。通过利用抗 6 $\times$ His 标签的小鼠单克隆抗体进行的蛋白质免疫印迹和质谱分析来确定 TAT-SALL4B 的重组蛋白(图 25C 和 25D)。关注 SALL4B 是因为它是一种形式较短并且在大肠杆菌中以高水平表达的蛋白质。TAT-SALL4B 处理后 3 天,CD34⁺ 细胞迅速地扩增(图 26A)。与 SALL4 融合蛋白(200nM)以及 TPO、SCF 和 Flt-3 配体一起培养 3 到 4 天的人 CD34⁺ 细胞显示出分别超过 10 倍和 8 倍的总单核细胞和 CD34⁺ 细胞增加(图 26B 和 26C)。一天添加两次 TAT-SALL4B (200nM),并且看起来足以使细胞扩增。此外,CFU 分析证实,这些细胞可以形成各种集落,包括 CFU-GM、CFU-GEMM 和 BFU-E (图 26D)。通过用经过 TAT-SALL4B 处理的细胞相对于对照物的增加倍数乘以 CFU 数目,注意到总体 CFU 数目增加约 10 倍。

[0262] 实例 10

[0263] 在 SF9 昆虫细胞中 TAT-SALL4B 融合蛋白的表达和纯化

[0264] 4 天的 SF9 细胞目视检测显示,在细胞内存在多角体蛋白(图 30),指示病毒产生。hSALL4B 蛋白的表达通过蛋白质免疫印迹得到确定(图 29)。这一印迹也确定了在来自 Ni-NTA 珠粒的第一洗出液和第二洗出液中存在标记 His 的 SALL4B。

[0265] 实例 11

[0266] 利用 TAT-SALL4B 蛋白在体外扩增全小鼠骨髓细胞

[0267] 如先前所述,分离小鼠全骨髓细胞。在 20ng/ml TAT-SALL4B 蛋白或 20ng/ml BSA 中培养细胞,以 0.4×10^5 个细胞开始。第 4、7 和 10 天,细胞数量在 BSA 组中平均为 1.16×10^5 个、 2.04×10^5 个和 3.16×10^5 个,并且在 SALL4B 组中为 1.89×10^5 个、 7.28×10^5 个和 10.1×10^5 个(图 31),表明 TAT-SALL4B 可以在体外促进全骨髓细胞增殖。

[0268] 实例 12

[0269] TAT-SALL4B 蛋白促进骨髓细胞扩增和骨髓恢复

[0270] 在致死性照射后 24 小时,连续 7 天将 TAT-SALL4B 蛋白、G-CSF 或 PBS 经腹膜内注射到小鼠体内。致死性照射的剂量(7Gy, γ 射线)能够在 1 周内杀死超过 99% 的小鼠骨髓细胞。通过冲洗一只一般类型小鼠中两侧的胫骨和股骨,可以获得平均 2×10^7 个全骨髓成核细胞。在 PBS 组中,照射后第 8 天,全骨髓细胞数量为 1.32×10^5 个($\pm 0.13 \times 10^5$; $n=6$)。比较而言, G-CSF 动物具有 4.51×10^5 个($\pm 0.43 \times 10^5$; $n=6$) 细胞,并且 SALL4B 处理组具有 7.91×10^5 个($\pm 0.75 \times 10^5$; $n=7$) 细胞。与先前的报导一致, G-CSF 可以使细胞数量增加 3.42 倍。在 SALL4B 组中相比 PBS 对照组和 G-CSF 组的增加倍数分别为 6.00 倍和 1.75 倍,表明 SALL4B 在加强照射后骨髓细胞增殖方面甚至优于 G-CSF。

[0271] 分析在照射后第 8 天从不同组得到的组织切片。与仅极少量细胞(主要是骨髓基

质细胞)存留在小鼠骨髓腔中的 PBS 组相比,骨髓的细胞化通过 G-CSF 或 SALL4B 处理得到急剧增强(图 32-34)。此外,大多数细胞都被鉴别为造血细胞。这些数据证实,SALL4B 有效地通过增加骨髓细胞增殖来促进骨髓恢复。

[0272] 实例 13

[0273] SALL4B 增加小鼠骨髓中造血干细胞和祖细胞扩增

[0274] 利用流式细胞术,检测针对全骨髓细胞具有 Lin、c-Kit 和 Scal-1 染色组合的造血干细胞和祖细胞的含量。与对照物(5.1%)相比较,HPC (Lin-/c-Kit+Scal-)的百分含量在 G-CSF 组中增加到 14.7% 并且在 SALL4B 组中增加到 9.82%。这些结果与 CFC (集落形成细胞)分析充分相关(图 35)。CFC 分析是得到良好表征的分析,可以检测造血祖细胞增殖频率对刺激剂或抑制剂起反应而发生的增加或减少。

[0275] 此外,与对照物(1.24%)相比较,造血干细胞(HSC) (Lin-/c-Kit+/Scal-1+) 百分含量在 G-CSF 组中也显著较高(图 36)。重要的是,SALL4B 组中的 HSC 百分含量甚至高于 G-CSF 组中的百分含量(约 2.7 倍增加)。在小鼠骨髓中 HSC 的总增加倍数(相对于对照物)在 G-CSF 和 SALL4B 中分别为 11.8 倍(n=6)和 32.2 倍(n=7)。

[0276] 实例 14

[0277] 利用 TAT-SALL4B 蛋白增强脐血干细胞和祖细胞的长期定植

[0278] 使用脐带血(UCB)细胞测试 TAT-SALL4B 增加骨髓移植效率的能力。将 20,000 个 CD34+UCB 细胞移植到经历亚致死照射的 NOD/SCID 小鼠中,并用 2 微克/天 SALL4B 蛋白治疗 7 天,随后每隔一天用 2 微克/天再治疗一周(图 37)。在移植后 16 周,用从 SF9 细胞分离的 TAT-SALL4B 蛋白处理的动物在外周血中 CD45+ 细胞(UCB 细胞)的长期定植方面显示相比用 PBS 处理的对照物的 8.5 倍增加。移植后 16 周的 CD45 细胞(UCB 细胞)与用 PBS 处理的对照组相比较。进一步检查移植于骨髓中的 UCB 细胞。如图 38 中所示,在授予 TAT-SALL4B 的小鼠体内供体细胞 UCB 的水平相比注射 PBS 到对照小鼠增加 10 倍,并且是在移植后 4 个月进行测量,证实了长期定植的实现。这些研究指出,TAT-SALL4 蛋白是一种促进干细胞定植的稳固因子。

[0279] 实例 15

[0280] TAT-SALL4 增加干细胞和祖细胞的产率

[0281] 使用造血性集落形成祖细胞(CFC)的计数来评价外周血祖细胞收集情况。CFC 与嗜中性粒细胞恢复之日的关联性(以相关系数测量)强于总核细胞的关联性。将 SALL4B 蛋白经腹膜内注射到野生型小鼠,随后从经过注射的小鼠收集外周血以评价 CFC 的数量。如图 39 中所示,来自注射 SALL4B 蛋白的小鼠的外周血中 CFC 的数量相比注射 PBS 的小鼠的外周血中 CFC 的数量增加三倍。这一研究指出,当授予 TAT-SALL4 时,TAT-SALL4 能够增加外周血的干细胞和祖细胞产率。

[0282] 讨论

[0283] 如本文中所证实,SALL4A 和 SALL4B 是造血干细胞扩增的较强正调控因子。尽管先前使用例如胎肝酪氨酸激酶(Flt3)配体、干细胞因子、白细胞介素 6 和 11、HOXB4、OCT4 和 Nanog 等造血生长因子扩增 HSC 的尝试仅显示有限的 HSC 扩增,但接受 SALL4A 或 SALL4B 表达的 HSC 能够实现高水平离体扩增。用 SALL4 转导的细胞的培养物引起广泛的 HSC 扩增,其扩增水平在 2 到 3 周内达到对照物的超过 1000 倍,并且扩增的 HSC 未显示或显示极少成

熟。而且,所述扩增的发生相当迅速,仅在数天内就出现显著 HSC 生长。此外, SALL4 诱导的 HSC 扩增未削弱造血细胞的分化。对于 HSC 来说, SALL4 看起来起到维持未分化的增殖状态的作用并且阻止细胞分化。

[0284] 博文描述了一种新的治疗策略,在一些实施例中,它使用了细胞因子依赖性 SALL4 技术使人 HSC 在 4 周内急剧离体扩增 10,000 到 15,000 倍,同时无显著分化。在一些实验中,甚至在细胞培养 8 周后,仍有 37% 的 CD34⁺ 细胞仍为 CD34⁺/CD38⁻ (数据未显示)。在异种移植模型中,经过 SALL4 离体诱导 4 周的细胞的干细胞频率仅为新鲜 CD34⁺ 细胞的 2 到 2.5 倍 (图 26H)。不希望受任何科学理论的束缚,用 SALL4 诱导的 HSC 的体内生长可能不同于在存在过量特定细胞因子的细胞培养条件 (离体) 中的生长。

[0285] 在本文中所述的一些实施例中, HSC 扩增的幅度极高,对于 CD34⁺/CD38⁻ 群和 CD34⁺/CD38⁺ 群来说为 10,000 倍。此外,利用本文中所述的 SALL4 方法进行的可定植长期 HSC 的扩增可以在本发明实施例中实现,并且得到了本文中所述的二次和三次移植研究的证据的支持。

[0286] CD34⁺ 细胞的大量离体扩增可以使用本文中所述的材料和方法实现,同时无分化。这些扩增的细胞保留了长期定植特性,与体内未经过操作的细胞类似。在本文中所述的实验实例中,通过连续异种移植模型和再群体化分析所证实,用 SALL4 扩增的细胞持续长期定植。最关键的是,在移植后,它们不会压倒小生境诱导的调控性控制,使这些扩增的干细胞能够避免白血病的形成。在经过连续异种移植的动物中或在移植后超过 10 个月,在经过移植的小鼠中没有明显白血病的迹象 (数据未显示)。此外,当 SALL4A 或 B 表达并引入小鼠造血干细胞 / 祖细胞中时,进行同系移植后 12 个月,未展现白血病的形成 (n=6)。

[0287] 参考文献

[0288] 1. 武威达 (Ueda) 等人,利用干细胞因子、Flk2/Flt3 配体、血小板生成素、IL-6 和可溶性 IL-6 受体扩增人 NOD/SCID 再群体化细胞 (Expansion of human NOD/SCID-repopulating cells by stem cell factor, Flk2/Flt3 ligand, thrombopoietin, IL-6, and soluble IL-6 receptor), 临床研究杂志 (J Clin Invest.) (2000) 105 (7):1013-21。

[0289] 2. 米村 (Yonemura) 等人,造血祖细胞的体外扩增和干细胞的维持:FLT3/FLK-2 配体与 KIT 配体之间的比较 (In Vitro Expansion of Hematopoietic Progenitors and Maintenance of Stem Cells: Comparison Between FLT3/FLK-2 Ligand and KIT Ligand), 血液学 (Blood) (1997) 89:1915-1921。

[0290] 3. 皮特斯 (Peters) 等人,利用白细胞介素 -3 (IL-3)、IL-6、IL-11 和干细胞因子进行鼠类骨髓细胞的离体扩增导致受照射宿主中定植减少 (Ex vivo expansion of murine marrow cells with interleukin-3 (IL-3), IL-6, IL-11, and stem cell factor leads to impaired engraftment in irradiated hosts). 血液学, (1996) 87 (1):30-7。

[0291] 4. 安托查克 (Antonchuk) 等人,HOXB4 诱导的成体造血干细胞的离体扩增 (HOXB4-induced expansion of adult hematopoietic stem cells ex vivo), 细胞 (Cell). 2002 109 (1):39-45。

[0292] 5. 艾琳 U (Elling U), 克拉森 C (Klasen C), 艾森伯格 T (Eisenberger T), 安拉格 K (Anlag K), 特雷尔 M (Treier M) (2006) 鼠类内细胞群来源的谱系取决于 Sall4 功能 (Murine inner cell mass-derived lineages depend on Sall4 function). 美国国

家科学院院刊(Proc Natl Acad Sci USA) 103(44):16319-24。

[0293] 6. 哈特 AH (Hart AH), 哈特雷 L (Hartley L), 易卜拉欣 M (Ibrahim M), 罗伯 L (Robb L) (2004) 小鼠和人体内促进多能性的 Nanog 基因的鉴别、克隆和表达分析 (Identification, cloning and expression analysis of the pluripotency promoting Nanog genes in mouse and human). 发育动力学 (Dev Dyn) 230(1):187-98。

[0294] 7. 张 J (Zhang J), 唐 WL (Tam WL), 童 GQ (Tong GQ), 吴 Q (Wu Q), 陈 HY (Chan HY), 等人 (2006) Sall4 通过转录调控 Pou5f1 来调节胚胎干细胞多能性和早期胚胎发育 (Sall4 modulates embryonic stem cell pluripotency and early embryonic development by the transcriptional regulation of Pou5f1). 自然—细胞生物学 (Nat Cell Biol) 8(10):1114-23。

[0295] 8. 李 SS (Li SS), 刘 YH (Liu YH), 曾 CN (Tseng CN), 丛 TL (Chung TL), 李 TY (Lee TY), 等人 (2006) 在台湾得到的五种新人类胚胎干细胞系的表征和基因表达谱 (Characterization and gene expression profiling of five new human embryonic stem cell lines derived in Taiwan). 干细胞与发育 (Stem Cells Dev) 15(4):532-55。

[0296] 9. 王 J (Wang J), 饶 S (Rao S), 楚 J (Chu J), 沈 X (Shen X), 勒瓦瑟 DN (Levasseur DN) 等人 (2006) 用于胚胎干细胞多能性的蛋白质相互作用网络 (A protein interaction network for pluripotency of embryonic stem cells). 自然 (Nature) 444(7117):364-8。

[0297] 10. 吴 Q (Wu Q), 陈 X (Chen X), 张 J (Zhang J), 乐 YH (Loh YH), 罗 TY (Low TY), 等人 (2006) 在胚胎干细胞中 Sall4 与 Nanog 相互作用并且共同占据 Nanog 基因组位点 (Sall4 interacts with Nanog and co-occupies Nanog genomic sites in embryonic stem cells). 生物化学杂志 (J Biol Chem) 281(34):24090-4。

[0298] 11. 周 Q (Zhou Q), 奇普菲尔德 H (Chipperfield H), 米尔顿 DA (Melton DA), 王 WH (Wong WH) (2007) 在小鼠胚胎干细胞中的基因调控网络 (A gene regulatory network in mouse embryonic stem cells). 美国国家科学院院刊, 104(42):16438-43。

[0299] 12. 陈 X (Chen X), 维嘉 VB (Vega VB), 恩格 HH (Ng HH) (2008) 在胚胎干细胞中的转录调控网络 (Transcriptional regulatory networks in embryonic stem cells). 冷泉港定量生物学讨论会会议录 (Cold Spring Harb Symp Quant Biol) 73:203-9。

[0300] 13. 林 CY (Lim CY), 唐 WL, 张 J (Zhang J), 安格 HS (Ang HS), 贾 H (Jia H), 等人 (2008) Sall4 在不同胚胎来源的干细胞谱系中调控不同转录环路 (Sall4 regulates distinct transcription circuitries in different blastocyst-derived stem cell lineages). 细胞干细胞 (Cell Stem Cell) 3(5):543-54。

[0301] 14. 杨 J (Yang J), 柴 L (Chai L), 福尔斯 TC (Fowles TC), 艾琳普 Z (Alipio Z), 许 D (Xu D), 等人 (2008) 全基因组分析揭露 Sall4 是鼠类胚胎干细胞中多能性的主要调控因子 (Genome-wide analysis reveals Sall4 to be a major regulator of pluripotency in murine-embryonic stem cells). 美国国家科学院院刊, 105(50):19756-61。

[0302] 15. 二木 (Futaki) 等人, 富精氨酸肽和其内化机制 (Arginine-rich peptides and their internalization mechanisms), 生物化学学会学报 (Biochem Soc Trans), 20074:784-7。

[0303] 16. 克罗尼坎(Cronican)等人,在体外和体内使用超动力蛋白将功能性蛋白有效递送到哺乳动物细胞中(Potent Delivery of Functional Proteins into Mammalian Cells in Vitro and in Vivo Using a Supercharged Protein),美国化学学会化学生物学(ACS Chem Biol).2010年6月21日。[电子版在印刷版前]

[0304] 17. 夏伯格(Sieburg)等人,造血干细胞隔室由有限数量的离散干细胞亚集组成(The haematopoietic stem cell compartment consists of a limited number of discrete stem cell subsets).血液学,2006107:2311-6。

[0305] 18. 斯查罗德 T (Schroeder T),造血干细胞异质性:亚型,并非不可预知的行为(Haematopoietic Stem Cell Heterogeneity:Subtypes,Not Unpredictable Behavior).干细胞,2010。

[0306] 19. 戴斯卡(Dykstra)等人,在体内不同造血细胞分化程序的长期传播(Long-Term Propagation of Distinct Hematopoietic Differentiation Programs In Vivo).细胞干细胞,20071(2):218-229。

[0001]

序列表

<110> 纽约州立大学研究基金会
 尤珀·玛 (Ma, Yupo)
 <120> 干细胞的离体扩增的方法和这样的细胞的用途
 <130> 2238/81894-A-PCT
 <150> US 61/376,122
 <151> 2010-08-23
 <160> 12
 <170> PatentIn version 3.3
 <210> 1
 <211> 5143
 <212> DNA
 <213> 智人 (Homo sapiens)
 <400> 1
 ttattctgcc ccagctgatg tttagaccag catgtcggg aggaagcaag cgaagcctca 60
 acatttccaa tccgaccccg aagtggcctc gctcccccgg cgagatggag acacagaaaa 120
 gggcacaacg agtcgcccta ctaagagcaa ggatgcccaac gtctgtggcc ggtgtgtgtg 180
 cgagttcttt gaattatcag atctcttctt ccacaagaag aactgtacta aaaatcaatt 240
 agttttaate gtaaatgaaa atccagcctc cccacccgaa acctctctcc ccagccccc 300
 tcttgataat cctgataaac aaatgaatga cacagttaac aaaaacagatc aagtggactg 360
 cagcgacctt tcagacaca acggaattga cagggaagag tccatggagg tggagccccc 420
 gggtgctaac aaaagggca ggggcacttc cagcggcagc cacagcagta cggccccaag 480
 cagcagcagc agcagcagca gcagcagcgg cggcggcggc agctcctcca caggtacctc 540
 agcgatcaca acctctctac ctcaactcgg ggacctgaca aactgggca acttctcgt 600
 aataacagc aacgtcatca tcgagaacct ccagagcacc aaggtggcgg tggcccagtt 660
 ctcccaggaa gcgagtgcg gggggcctc tgggggcaag ctggcgttc cagccctcat 720
 ggaacaactc ctactctgc agcagcagca gatccaccag ctgaattga tcgaacagat 780
 tggcaccaaa atattgcgtg tggctctca gaatgcagac ttgccaacat ctctatgtc 840
 ttctcaaggt actttacgaa catctgcaa ccccttgctc acgctaagtt cccatttacc 900
 tcagcagctg gcagcagcag ctggatttgc acagagcctc gccagccaat ctgccagcat 960
 tagtggtgtg aaacagctac ccccaatcca gctacctcag agcagttctg gcaacacct 1020
 cattecatcc aacagcggtt cttctcccaa tatgaacata ttggcagcgg cagttaccac 1080
 cccgtctctt gaaaaagtgg ctccaagtgc tggggcctcc catgtcagca acccagcggt 1140
 ctcatcatcg tctcaccag ctttgcgaat aagcagttta ttaagtcctg cgtctaatcc 1200
 actctacat cagcaagcct ccgctaactc gggtttcccc agcccttgc ccaacatcgg 1260
 aacaactgca gaggatttaa actccttgc tgccttggcc cagcaangaa aaagcaagcc 1320
 accaatgtc actgcctttg aagcgaaaag tacttccgat gaggcattct tcaaacacaa 1380
 gtgcaggttc tgcgcgaagg tctttgggag tgacagtgc ttgcagatcc acttgcgttc 1440
 ccataccgga gagaggccat tcaagtcaa catctgcggg aacaggttct ccaccaaggg 1500
 gaatctgaaa gtccacttcc agcgccacaa agagaaatac cctcatatcc agatgaacce 1560
 ctatctgtg cctgagcatt tggacaatat ccccaagagt actggcatcc catatggcat 1620
 gtccatccct ccagagaagc cagtcaccag ctggctagac accaaaccag tctgtcctac 1680
 tctgaccact tcagtcggcc tgcgttgc cccaaccctc ccaagcctca tacccttcat 1740

[0002]

```

caagaaggaa gagccagccc ccattcccatt cagccattct gccaccagcc ccccaggctc 1800
agtcacaaagt gactccgggg gccctgagtc agccacaaga aacctagggt ggctcccaga 1860
ggaagccgaa gggtcacac tgcacacctc tggtaggcaa agcgaagaga gtgcatgtgt 1920
caccacactca gtcccgacgg cgagcagtag cgtcctgagc tcccagcggg cagaactggg 1980
ccccgggggc agtgcacaca ccttcacaaa cctttgtgtt cgcctcatgt ccgagcagtt 2040
caaggccaaag tttcttttg ggggacacct ggaclcagcl caggcatcag agacgtccaa 2100
gttcacgcaa ctggtagaaa acattgacaa gaaggccact gaccccaatg agtgcatcat 2160
ctgccaccgg gttctcagct gccagagcgc cttgaaaatg cactacagga cacacactgg 2220
ggagaggccc ttttaagtga agatctgttg ccgggctttc accacgaaag ggaattctaa 2280
aaccacactac agtgtccatc gtgctatgcc cccgctcaga gtccagcatt cctgcccatt 2340
ctgccagaag aagttcacga acgtgttgtt cctgcagcag cacatccgaa tgcattatggg 2400
agccacagat cccaacaccc cagtcaccga cagctactct gagtccatgg agtgtgacac 2460
aggttccttt gatgagaaaa attttgatga cctagacaae ttctctgatg aaaacatgga 2520
agaactgtct gagggcagca tccctgatac acctaaagtct gcagacgcct cccaagacag 2580
cttatccclt tcgcctttgc ccttcagatc gtgcagcacc gtgtctttgg aaatcagat 2640
gaagatgac aatgtctggc tggcagagca gctacagccc agcctgaagt cagtggagaa 2700
tggtgtccatc gagggggatg tcttgaccaa tgaattcatc tcagtgggtg gtgacatgga 2760
aagccaaagt gctggcagcc cagccatctc agagtctacc tcttccatgc aggcctctgt 2820
cccgccaac agcacgcagg agttccacaa gtcacccagc attgaggaga aaccacagag 2880
agcgttccca agcagatttg ccaatggttt gtctccccc ccagtgaatg gtgggcttt 2940
ggatttgaca tctagtcacg cagagaaaat cateaaagaa gattcttttg ggatectctt 3000
ccccctttaga gaccggggta aatttaaaaa cactgcttgt gacatttgtg gcaaaacatt 3060
tgtttgtcag agtgctttgg acattcacta tagaagtcac accaaagaga gacctttat 3120
ttgcacagtt tgcacatctg gcttttccac aaagggtaat ttgaagcagc acatgttgac 3180
acatcagatg cgagatctgc caccacagcl ctltgagccc agttccaaac ttgcccacaa 3240
tcagaactca gcggtgatcc ccgccaacte gttgtcatct ctcateaaga cagaggtcaa 3300
cggtctctgt catgtttctc ctccagacag taaggacacc cccaccagtc acgtcccgtc 3360
tggtctctgt tcttctctgt ccacatcccc agttctgtct cctgtctctg ccaggagaa 3420
tcccagcag cactactgca acacatgttg caaaaccttc tctctatgca gtgcctgca 3480
gattcagcag agaactcaca ctggagagaa accctttgct tgcactatct gtggaagagc 3540
tttcacgacl aaaggcaatc ttaaggtaca catgggcaact cacatgtgga atagcaccac 3600
tgcaagcagg ggtcggcgcc tctctgtgga tggccccatg acatttctag gaggaatcc 3660
cgtcaagttc ccagaaatgt tccagaagga tttagcgcca agatcaggaa gtggggatcc 3720
ttccagcttc tggaaatcag atgcagcagc gctctccaac gggttgccga tgaaggccaa 3780
cgagatctcc gtcatcaga acggtggcat cctccaaatt cctggaagcc tggcagtggt 3840
gaacagctca cctgttagtg ggtgacggg aaacctggag aggtccaga actcagagcc 3900
caatgtctcc ctggccggcc tggagaaaaa ggcaagcagt gagaacgaa ccaacttccg 3960
cttcacccgc ttctgtggag acagcaagga gatcgtcacg agttaagca gctcgggctg 4020
gagacatage attcattcct gttcagaatg cgacctatgg tggcctccca ctcttgcgc 4080

```

[0003]

```

cccacccccc cccgccctt ccttctgttc cccagatcta tgaactacaa cattatgaag 4140
acattctttt gtaccttggt caactitaga gttetaagaa agcttattta ttagcgatat 4200
aaccttgcct tgcacacaga atgcaagcgt taacttttgt cttctgtatt ttggactaaa 4260
tactaatga ctagagtgtc gtaaacctgc tgtaacattt atggcaattg caagttgccc 4320
tgcctaggcag ttgtaatttg gcattaaact atttctata tccagtttaa tatgaattcg 4380
gtgttgatgc aatgctcag tgatgcatta gatctctaat aaagtctgta tatacatgia 4440
cactttgate ctgctggaaa attttatcag caaacacatt gtctaacttt tcaaacaga 4500
tttaaggaaa ggactgaaag tacagactga acagtgttgt tcttgaaag gtttggtttt 4560
ttaatttita ttctaaaatt caaccttttt tttgtcgat ttaaccattt ccattttgaa 4620
ctgctatttg tatgtgctt ttaacttgag tegtcttcaa tgttaataag ttctgtaca 4680
gtaataanga cgcagaattc tttagagaaa aagaaaacaa gcgttggttt ggtagtga 4740
actgagacgt aacattttgc cttgtaggta tattcacgat agaaaatgtg tcttggaatt 4800
tcacaatget gctaatgafa gcattttgaa caaccttcag tggagaaaaa gtagatgc 4860
ttgtatatac aataangaat atcactttca ttcaaatgta catatgttcc ttacaagagc 4920
aaatgcttct tcttgatcaa gagagcaggt atagtgtttg tttattttgt cttaggatg 4980
gaagaaaaaa attggactgt tacatgcact ttcttgaaa gttgaaagga aaggggggt 5040
ccaatttctt taacatttaa tacttactaa caacagagat acgttaattt tactcaagta 5100
atcaaataca ttttttttgc aacagataaa acaaaatact gtg 5143

```

<210> 2
 <211> 5253
 <212> DNA
 <213> 智人(Homo sapiens)

```

<400> 2
ctttggaggt tcagagctgc aagaagatgg ggactgggtg cggggcgcca gcgcctgacc 60
cgtctgggtt gcggccgggg tggagagtgc tgcctggcgc cagtgttcc ggagacggtg 120
caaacggacg gggaaagtgt cgggtcttgg ctgcacaaat ttatctccgc atctcttccc 180
acaacacttg cacctctgc ccccaaaat cttctggag acacagaaaa gggtaaccg 240
agtcgcctca ctaagagcaa ggatgccac gtctgtggcc ggtgctgtgc cgagtcttt 300
gaattatcag atctctctgt ccacaagaag aactgtacta aaaatcaatt agttttaatc 360
gtaaatgaaa atccagctc cccacccgaa acctctccc ccagccccc tctgataat 420
cctgatgaac aaatgaatga cacagttaac aaaacagatc aagtggactg cagcgacct 480
tcagaacaca accgaattga cagggaagag tccatggagg tggaggcccc gtttgcatac 540
aaaagcggca gcggcactic cagcggcagc cacagcagta ccgcccacag cagcagcagc 600
agcagcagca gcagcagcgg cggcggcggc agctcttcca caggtacctc agcgtacaca 660
acctctctac ctcaactcgg ggacctgaca acactgggca acctctccgt aatcaacagc 720
aacgtcatca tcgagaacct ccagagcacc aaggtggcgg tggcccagtt ctcccaggaa 780
gcgaggtgcg gcggggcctc tgggggcaag ctggccgtcc cagccctcat ggaacaactc 840
ctagctctgc agcagcagca gatccaccag ctgcaattga tcgaacagat tctgcacca 900
atatgtctgt tggcttctca gaatgcagac ttgccaacat cttctagtc ttctcaaggt 960
actttacgaa catctgcaaa cccctgtcc acgctaagtt cccatttate tcagcagctg 1020
gcagcagcag ctggatttgc acagagctc gccagccaat ctgccagcat tagtggtgtg 1080

```

[0004]

aaacagctac ccccaatcca gctacctcag agcagttctg gcaacacct cattecatcc	1140
aacagcggtt cttctcccaa tatgaacata ttggcagcgg cagttaccac cccgtccctt	1200
gaaaaagtgg ctccaagtgc tggggcctcc catgtcagca acccagcggc ctcatcatcg	1260
tctcaccag cttttgcaat aagcagtta ttaagtcttg cgtctaatcc attctacct	1320
cagcaagcct ccgctaacct gggtttcccc agccctttgc ccaacatcgg aacaactgca	1380
gaggatttaa actccttgtc tgcctlggcc cagcaaagaa aaageaagcc aocaaatgic	1440
actgcctttg aagcgaaaag tacttccgat gaggcattct tcaaacacaa gtgcaggttc	1500
tgcgcgaagg tctttgggag tgacagtgc cttgcagatcc attgcgttc ccataccgga	1560
gagaggccat tcaagtgcac catctgcggg aacaggttct ccaccaaggg gaatctgaaa	1620
gtccacttcc agcgcacaaa agagaataac cctcatatcc agatgaaccc ctatctctgt	1680
cctgagcatt tggacaalat ccccacgagt actggcatcc catatggcat gttccatcc	1740
ccagagnagc cagtcaccag ctggctagac accaaaccag tcttgctac ttgaccact	1800
tcagtcggcc tgcgttgcc cccaacctcc ccaagctca tacccttcat caagacggaa	1860
gagccagccc ccctcccat cagccattct gccaccagcc ccccagctc agtcaaaagt	1920
gactccgggg gccctgagtc agccacaaga aacctaggtg ggtccccaga ggaagccgaa	1980
gggtccactc tgcaccctc tgggtgcaaa agcgaagaga gtggcatggt caccaactca	2040
gtcccgacgg cgagcagtag cgtcctgagc tcccagcgg cagactggcg ccccgccggc	2100
agtgcacca ccttcacaaa cctttgttg ccgtctatgt ccgagcagtt caaggccaag	2160
tttctttttg ggggactcct ggactcagct caggcatcag agcgtccaa gcttcagcaa	2220
ctggtagaaa acattgacaa gaagccact gaccccaatg agtgcactat ctgccaccgg	2280
gttctcagct gccagagcgc ctgaaaatg cactacagza cacacactgg ggagagccce	2340
tttaagtgtc agatctgigg ccgggttttc accacgaaag ggaattttta aaccactac	2400
agtgtccatc gtctatgcc ccgctcaga gtccagcatt cctgcccat ctgcagaaag	2460
aagtccaaga acgtgttgtt cctgcagcag caccatcgaa tgcataatgg aggccagatc	2520
cccaacaccc cagtcccga cagctactct gactccatgg agtctgacac aggttccctt	2580
gatgagaaaa attttgatga cctagacaa ttctctgatg aaaacatgga agactgtcct	2640
gagggcagca tccctgatac acctagtct gcagacgctt cccaagacag ctatctctt	2700
tcgcctttgc ccctcgagat gtgcagatc gctgctttgg aaaatcagat gaagatgac	2760
aatgttggtc ttgcagagca gtacaggcc agccigaagt cagtggagaa tgggtccatc	2820
gaggggatg tctgaccaa tgattcatcc tcagtgggtg gtgacatgga aagccaaagt	2880
gttgccagcc cagccatctc agagtctacc tcttccatgc aggtctgtc cccgtccaa	2940
agcagcagg agttccacaa gtccaccagc attgaggaga aaccacagag agcggtccca	3000
agcgagtttg ccaatggttt gtctccacc ccagtgaatg gtgggcttt ggatttgaca	3060
tctagtccg cagagaaaat catcaagaa gattctttgg ggatctcttt cctttttaga	3120
gaccggggta aatttaaaaa cactgtttgt gacattttgt gcaaaacatt tgcctgtcag	3180
agtgccttgg acattcacta tagaagtcac accaaagaga gaccatttat ttgcacagtt	3240
tgcaatctgt gcttttccac aaagggtaat ttgaagcagc acatgttgac acatcagatg	3300
cgagatctgc cctccagctt ctttgagccc agttccaaac ttggcccca tcagaactca	3360
gcgtgtatc ccgccaactc gttgtcatct ctcatcaaga cagaggctca cggtttcgtg	3420
catgtttctc ctccaggacg taaggacacc cccaccagtc acgtccctgc tgggcctctg	3480

[0005]

```

tcttctcttg ccacatcccc agttctgtct cctgtctctg ccaggagAAC tcccaagcag 3540
cactacttga acacatgttg caaaaccttc tctctataga gtgcccctga gattcagcag 3600
agaactcaca ctggagagaa acccttttgt tgcactatit gtggaagagc tttaacgaat 3660
aaaggcaatc ttaaggtaca catgggcact cacatgtgga atagcacccc tgcacgaagg 3720
ggteggcgge tctctgtgga tggccccatg acatttctag gaggaatcc cgtaagttc 3780
ccagaaatgt tccagaagga ttggcgga agatcaggaa gtggggatcc ttccagcttc 3840
tggaaatcag atgcagcagc gctctccaac gggctggcga tgaaggccaa cgagatctcc 3900
gtcattcaga acggttgcct cctccaatt cctggaagcc tggcagtgga gaacagctca 3960
cctgttagtg ggctgacggg aaacctggag aggcctcaga actcagagcc caatgctccc 4020
ctggccggcc tggagaaaat ggcaagcagt gagaacggaa ccaacttccg ctccaccgcc 4080
ttcgtggagg acagcaagga gatgttcacg agttaagca gctcgggctg gagacataga 4140
attcattcct gtccagaatg cgacctatgg tggcctctca ctcttgcgc cccaccgcc 4200
cccgcctctt ccttctgttc cccagatcta tgaactacaa cattatgaag acattctttt 4260
gtaccttggt caactttaga gtcttaagaa agcttattta tttagcatat aaccttgctt 4320
tgcaaacaga atgcaagcgt taactttgtt ctctgtat ttggactaaa tactaatga 4380
ctagagtgtt gtaaacctgc tgaacattt atggcaattg caagttgcc tgetaggcag 4440
ttgtaatctg gcattaactt attttctata tccagtttaa tatgaatctg gtgttgatgc 4500
aatgcctcag tgaatcatta gatctctaat aaagtctgta tatacatgta cactttgata 4560
ctgetggaaa attttatcag caaacacatt gtctaattt tcaaacaga tttaaggaaa 4620
ggaotgaaag tacagactga acagtgtggt tcttgaag gtttggtttt ttaattttt 4680
ttctaaaatt caaccttttt tttgtctgat ttaaccttt ccattttga ctgetatttg 4740
tattgtgctt ttacttgag tctgttcaa tgttaataag ttctgtaca gtaataagca 4800
cgcagaattc tttagagaaa aagaaaacaa gcttgtttt ggtagttaga actgagacgt 4860
aacattttgc ctgtaggta tattcagat agaaaatgtg tgcgtgaatt tcacaatget 4920
gctaagata gcatcttag caaccttcag tggagaaaat gtagatctc ttgtatatac 4980
aataagaaat atcactttca ttcaaatgta catatgttcc ttacaagagc aaatgcttct 5040
tcttgatcaa gagagcaggt atagtgttg tttattttgt cttaggtatg gaagaaaaa 5100
attggactgt tacatgcact ttcttgaaa gttgaaagga aagggggggt ccaatttctt 5160
taacatttaa tactactaa caacagagat actgtaattt tactcaagta atcaataca 5220
ttttttttgc aacagataaa acaaaatact gtg 5253

```

<210> 3

<211> 4931

<212> DNA

<213> 智人 (Homo sapiens)

<400> 3

```

gagctgcaga agcgtaggga agaagctgaa gaaaaaaagg gggcgtctcc cctttaaaga 60
cttgcaaga ttgagagaga aagagagaga gtaagaaca gagaatcaga gagagagaga 120
gagtctgtgt ctctgggaaa gaagaacatc tctgcttca acgtatttgc gctgggggag 180
aggcatcaat tggcttcgga cccaaggggg agacgagacc aggtcacccc ggttaagacc 240
aagtgagcgt tgcccccc cctctccaaact ctctaccggg gaatgtctcg gcgaaagcag 300
cggaaacccc aacagttaat ctgggaactg gaaggtecca ggcgtctga gnaeggtgat 360

```

[0006]

gclagcgagg	aggatcacc	ccaagtctgt	gccaaagtgt	ggcacaait	cactgaccca	420
actgaattcc	tgcgccacca	gaacgcatgt	ctactgacc	ctctgtaat	ggigataatt	480
gggggceagg	agaaceccaa	caactcttgc	gcctctcttg	aaceccggcc	tgagggtcac	540
aataatectc	aggteatgga	cacagagcat	agcaaceccc	cagattctgg	gtccctccgtg	600
cccacggatc	ccacctgggg	ccagagagg	agaggagagg	agtctccagg	gcatttccctg	660
gtcgctgcca	caggtaacgc	ggctggggga	ggcgggggce	tgatcttgcc	cagtcaccaag	720
ctgggagcaa	ccccattacc	tccagaatcg	acccctgcac	ccccctctcc	tccaccaccc	780
ctctccccc	caggggtagg	cagtggccac	ttgaatatcc	ccctgatctt	ggaagageta	840
cgggtgtctc	agcagcgcca	gacctcatcg	atgcagatga	ctgagcaaat	ctgcaggcag	900
gtctgtttgc	tiggtctctt	aggccagacg	gtgggtgccc	ctgccagtcc	ctcagageta	960
ccctggacag	ggactgccc	tccaccaag	ccccactac	ccctcttcag	ccccatcaag	1020
cctgtccaaa	ccagcaagac	actggcatct	tctctctct	ctctctcttc	ctcttcaggc	1080
gcagaaacgc	ccaagcaggc	ctcttccac	ccttaccacc	cactgggggc	acagcatctt	1140
ttctctgtctg	gaggggttgg	gcgaagccac	aaaccacccc	ctgccccttc	cccagccttg	1200
ccagccagca	cagatcaget	gattgcctcg	ctctatcttg	catttcccaag	caccacggga	1260
ctaetggcag	cacagtgtct	tggggcagcc	cgaggccctg	agccactgc	ctccccaggc	1320
ctctgaagc	caaagaatgg	aagtgggtgag	ctgagctacg	gagaagtgt	gggtcccttg	1380
gagaagcctg	gtggaagcca	caaatgcgcg	ttctgtgcca	aagtatttgg	cagtacacgt	1440
gccccgcaga	tcaactctg	tcccacacg	ggtgagaggc	ctataagtg	caatgtctgt	1500
ggaaacccgt	ttaccacccg	tggcaacctc	aaagtgcatt	tcaaccggca	tcgtgagaag	1560
taccacatg	tgcagatgaa	cccacacca	gtaccagagc	acctagacta	tgtcattacc	1620
agcagtggct	tgccttatgg	tatgtccgtg	ccaccagaga	agcccgagga	ggagccagcc	1680
actccagctg	gaggggttga	gcgaagcct	ctggtggcct	ccacaacagc	actcagtgcc	1740
acagagagcc	tgactctgt	ctccaccagt	gcaggccagc	ccacggctcc	aggactccct	1800
gctttcaata	agtttgtgt	catgaaagca	gtggaacca	agaatnaagc	tgtgaaaaac	1860
acccccccag	ggagtggagg	ctcagccatc	agtggagtgg	cagaaagtag	cacggcaact	1920
cgcattgcaac	taagttaagt	ggtgacttca	ctaccaagct	gggcactgt	taccaaccac	1980
tccaagtcca	ctggcagctt	ccccctccc	tatgtctag	agcccttggg	ggcctcacc	2040
tctgagacat	caaagctgca	gcaactggta	gaaaagattg	accggcaagg	agctgtggcg	2100
gtgacctcag	ctgcctcagg	agccccacc	acctctgcc	ctgcaccttc	atcctcagcc	2160
tcttttgga	ctaaccagt	tgtcatctgt	ctcagagtgc	ttagctgtcc	tggggcccta	2220
cgccttcatt	atggccaaca	tggaggtgag	agcccttcca	aatgcaaat	gtgtggcaga	2280
gccttctcca	ccagggttaa	ctcgcgtgca	catttctggg	gccacaagge	cagtcagct	2340
gccccggcac	agaattctctg	ccccactgc	cagaagaagt	tcaccaatgc	tgtactcttg	2400
cagcagcatg	tccgatgca	ctgggggggc	cagatcccca	acggtggtag	tgcactccct	2460
gaagggtggag	gagctgtcca	ggagaatggc	tccagcaat	ctacagtctc	cggggcaggc	2520
agtttcccc	agcagcagtc	ccagcagcca	tcaccggaag	aggagtgtgc	tgaggaggag	2580
gaagaggagg	ataggaaga	agaggaagat	gtgactgatg	aagattccct	ggcaggggaga	2640
ggctcagaga	gtggaggtag	gaaggcaata	tcagtggag	gtgattcaga	agaggcatct	2700
ggggcagagg	aggaggtggg	gacagtggcg	gcagcagcca	cagctgggaa	ggagatggac	2760

[0007]

```

agtaatgaga aaactactca acagtcctct ttgccaccac caccaccacc tgacagcctg 2820
gatcagcctc agccaatgga gcagggaagc agtgggtgtt taggaggcaa ggaagagggg 2880
ggcaaacccg agagaagctc aagtcgggca tcagcactca cccagaagg ggaagccacc 2940
agegtgacct tggtagagga gctgagcctg caggaggcaa tgagaaagga gccagagag 3000
agcagcagca gaaaggcctg cgaagtgtg ggcagcctt tccctccca ggcagctctg 3060
gaggagcate agaagacca cccaaggag ggcgcctct tcactgtgt ttctgcagg 3120
caggcttctc ttgagcggc taccctcaag aagcatatgc tctggcaca ccaccaggta 3180
cagccctttg ccccatgg cctcagaat attctctctc ttctctagt cctggctgt 3240
tcgcttcca tcacctccac agggtctcc cctttcccc gaaaagatga cccaagatc 3300
ccatgagcct gttttctgt acctgtgtc cttgtccca cagagcagaa acagcttcac 3360
aaaaggacct ccagagttc tgagccctga tttgtcttt ttctctaagt tcttaacatg 3420
ttatgtccct agtggcttt ctgtagtccc tgagcttggg aattactgtg cttacaaggg 3480
gatggccccc taaggaaatt ttctccctc ctctctctt gtaetgagg aacatagatt 3540
ctctcagct ttctcaagg gaacctctc cagcttccct ggtgtgacce ttctccccc 3600
tcctctctcc tctcccttc ccttggtag gtgcacctga gcacctacat ttggcattgc 3660
agcctagcca aaaaggctg gcagctgtc ctggagggcc cagtgcact cctctgggt 3720
gacctttctg ctacagctgt gggatgggt cccctatctc tctagaacca gtatgtggca 3780
ttctgtcaa atgacctgc catgaagccc tggaattcca gctccacct cactaccact 3840
ccaagcctgg cccaccagt gctgtttggc ctaggaaatg tggtgggaa ggtgcctcca 3900
acaatgggat ccagggaagc caaggagaag acagccccc tctatttcca gctctctga 3960
cccaaggcag tgcctagaa gccatcata gacaagaagt agcaaacgt acattcctc 4020
ttctccccc tgcctcagaa ggtgccgta ctgaagatgc tccagtaatt ggtgacccaa 4080
ccttaggaag tagggagaaa tgaaggagg gcataggaaa attttccag taaatccct 4140
gatggtcaca ttaagtaaa ggttttggc ggtcagtgtg ccaagacctc tccagcttct 4200
cattcatgat gacctctcaa agttgggaaa caagctgatt tcttgcaag aggtctccca 4260
ggagatattt gggaaatgt aagtctgtc cttaaggag ctttttggc cagcatgggt 4320
gatgaactaa tgaigagaga gtaaggaat gttctagaa catagggett gcfgtaacct 4380
atgtgactaa gaaaggaca tgatgaagg gaaaaggcct caattcttg tgaatgtgga 4440
cattctctgt aatattctt tgggctaata gtgacatagt gtgcagaggt gtaccaggga 4500
tcatggggga ttctctagca ctgatatgt tctagtttta gataactccc tctttatc 4560
cctggccctc tctatttcc ttatcttct cttcaagac cctaccctt ttgctatc 4620
cgtaggctgg gcttgtgtc ttgtcattg tctgttctt aagagtccta gctcagggt 4680
gctctctccc tgcctctccg tctgtaatg agttgtatg ttactctta acataggatc 4740
atttgaaca ggagttctg ggaggagag gtgagggtt tctattgac tgaattgaac 4800
gatgcttct cctcaagct taggtccag agcttctaa cctagtaaaa tgtcaagaac 4860
agacgggaga tattagtgc ttccctcta tcattaaagg tgttttaacc aaaaaaaaa 4920
aaaaaaaaa a 4931

```

<210> 4
 <211> 4775
 <212> DNA

[0008]

<213> 智人(Homo sapiens)

<400> 4

atgtctcggc gcaagcagge caagccccag caactcaagt cggacgagga gtgtctgccc	60
cctgaegggg cccccagca cgcgcgcccg ggggaagggt cggaggacgc agacagcggg	120
cccagagacc gcagcggggg cgaggagacc agcgtgtggc agaaatgcgt cgcagattc	180
ttaagtggtg cggacttctt ggagcaccag cggagctgca ccaagctccc gcccggtgtg	240
atgtgtcagc aggcagcgcg cgcgcgcccg cccgaggact tcccagacc ttcgcccgc	300
agctccccc cgcagcgcgc gaaagcgag gcggccgagg aggcgggtgc ggaggcgcgc	360
gaggcagagg ccagccgggt ggagaaggag gccgagccca tggacgcgga acccgcgggg	420
gacacgcgcg cgcgcgggc cccgctgtgc gccctgcac ccccaagccc cgcctacggc	480
gcgcccagca ccaacgtgac cctggaggcg ctgtctgagc ccaagggtgc ggtggcgcag	540
ttctcgcagg gcgcgcgcgc gccagcggcg tcgggagcag gtggaggcgt gccagctgca	600
gccgtgccc tgaacctgga acagctcatg gccctgcagc agcagcagat ccaccagctg	660
cagctcatcg agcagatccg cagccaggtg gccctcatgc agcgcgcgc gcccgggccc	720
tcactcagcc ccgcgccgc cccgagcgca ccgggcgcgg ccccagcca gtgtccnggg	780
ctggcgcgc tcccgctgtc gcccggggc cctgcgcgc ccatcgccgg ctggggccc	840
gccgcccgcg ccgccttgcg ggccgcgcag ccgtgtccc gccccagtc tggcgcagc	900
acccccggcg gccctgcgga gcccgagcgc ccgcgcgc cccagccgc cctgtcccc	960
gtgtccccc ccccgccgc agcgcgcag agcgcagcct cgtgcagcc gcagagcgca	1020
tccacgcgc ctgccttgc cccgggtccc ctgttggtg cggcgcgcg cctgccaagt	1080
ccgtttctac ctcagacttc cgcagcggc gtcattcttc ccaaccgcgt ggtcagctc	1140
gggccaagc ccaacgctct ggaccgcgt tcgcgcctca tgaagaccg caaggccaag	1200
ccgccaatg tgtcgtgtt cgagcccaaa gccagcgcg aggaaccgtt ctccaagcac	1260
aaatgccgt tctgcgcga ggtcttgcg agcgcagcgc cgtccagat ccacctgcgc	1320
tgcacacag gcgagcggc ctccaagtgc aacatctgcg ggaaccgctt ctccacaaa	1380
ggcaacctga aggtgtcaatt ccagagcgac aaggagaagt accccacat ccagatgaac	1440
ccttaccgcg tcccagata cctggacaa gtgcccacct gctgggcat cccctaagc	1500
atgtctgtgc cccccagaa gccctgacc acctggctgc acagcagcc cgtgtgtccc	1560
acctgtccc cgtcgtggg gtgtcaactg ccgcccactg tccctggcgc gcacggctac	1620
gccgactct cccagccac cccagccagc cgtccccc agagccctc gccgcctcc	1680
agcagtgct cctccttgc cccagcctc aaccaagtgg agtccggcgt gtggccccc	1740
gccgagtc cccagtcgt cctggcggg ccgcctctca ctaaagccga gcccgctagc	1800
ctgcccctga ccaacgcag gcccggggac gctcccgtgg gcgcgcagc tagcgtgca	1860
cccacatcgg tggacggcgc acccagcagc ctgggcagcc cggggtgtcc cgcgtctcc	1920
gagcagttca aggcacagtt tccgttcggg gggtgtctag actcgatgca aacgtggaa	1980
acctgaagc tgcagcagct ggtggagaac atcgacaaga agatgacgga cccgaaccag	2040
tgcgtcatct gccaccgggt gtgtgagctgc cagagcgcgc tgaagatgca ctaccggacg	2100
cacacggggg agcggccgtt caagtgcag atctgggcgc gcccttcac caccagggc	2160
aacctcaaga cgcacttgcg cgtgcaccgt gcaaaagccc cctgtcgcgt gcagcactcc	2220
tgccecatct gccagaagaa gttcaccaac gccgtgttcc tgcagcagca catccgcatg	2280

[0009]

cacatgggcg gccagatccc caacacgcgc ctgcggagg gettcagga tgcacggac	2340
tcagagctgg cctacgacga caagaacgcg gagacctga gcagctacga tgacgacatg	2400
gaagagaact ccatggagga cgaagctgag ctgaaggacg cggccacga cccggccaag	2460
ccactcctgt cctacgggg gtctgcccgc cctcccgcgc cctcggcat ctccagcatt	2520
gcgcgcctgg agaaccagat gaagatgac gacccggta tgagctgcca gcagctgacc	2580
ggctcgaagt ccgtggagaa cgggtccggg gagagtgac gccagagaa cgaactctcg	2640
tcggccgtgg gcgacctgga gagccgcgc gcgggcagcc cgcacctgc cgaactctcg	2700
tcctgcaggg cctgtgcgc ggcgccgcgc aatggtgaga gcttcgcgc caagtcgcgc	2760
ggcctggggc ccccgaggga gcccaggaa atcccgctca agaccagag gccggacagc	2820
ccagccgcgc ccccgggcag cggagcgcc cctggccgc gcggcaccac ggaggagggc	2880
cccttcagcc tctgttctc gagcagggag cggggtaagt gtcccagcac tctgtgtgtg	2940
gtctgtggca agccttttgc ttgcaagagc gcgttggaat tccactacc cagccatact	3000
aaggagcggc cattcgtctg cgcctctgc aggcgaggg gtccactat gggtaatita	3060
aaacagacat taatgacaca cagattgaaa gactgcctt ctcaattatt tgaccccaac	3120
tttgccttag gtccagcca aagcactct agcctgatct ccagcgccgc acccaccatg	3180
atcaaaatgg aagtgaacgg tcacggcaag gccatgggc tggcgaggg tccccgcctg	3240
cccgccggcg tccaggtccc cgcgggcct cagacagtga tgggcgggg cctggcgccc	3300
atgctggccc cccacccgc ccggacgccc aagcagcaca actgcagtc gtgcgggaag	3360
acctctctc cggccagcgc cctgcagatc catgagcga cgcacaccgc cgagaagccg	3420
ttcgctgca ccatctggc cgggccttc accactaagg gcaacctca gggtcacatg	3480
gggacacaca tctggaataa cgcgccgcg agacgcggc gcgcctgtc tctggagaac	3540
cccatggctc tccatgggg tgatgcctg aagtctctg aaatgttcca gaaggacctg	3600
gcagctggg caatgaacgt cgaaccagt ttttgaacc agtatctgc agccatcact	3660
aacgggtctg ccatgaagaa caacgagat tccgcatcc agaacggcg caiccccag	3720
ctcccgtga gtcttgggg cagcgccct cccctcttg gcagcatgc cagtggtatg	3780
gacaaagcac gcaatggag tagccaccc atctcagct tggacaaag cagctcagaa	3840
acagcagcca gccgccatt cagcggttt atcgaggata acaaggagat tggatcaac	3900
tagccagtga ctgcctcatc tgcctgcgc aggcaccagt ttgaagttg gagcatcagg	3960
cctccagact ttcttgctc ggtctcatt acacttcaa ccatagcaga aaacacttg	4020
tgcggctgcc gagagggtg ctgtgaagc ctgcctggc ctccctcaa cagcaagcct	4080
gactgtctc gagaactct caatcttta aataagctt ctcaaaaaa aaaagtgtt	4140
ggaaaaccgc cttaggaaca gaaagagct agaccatgc cacttcttt ctctgaaac	4200
ctaataatct ctccgagga gaaagggtt ctctcggtt tccagtgaa acatcttga	4260
tgtttcttt tgaattagt agacactga accgtgttt ttagaactct tcatgttaa	4320
gacgtgttt agtactccca atctgtgtg tcatgacaat atctctgtc gtatattta	4380
tgatgttaag ataatgggg taacagacaa tataatagc ccgacctaa accaagcttt	4440
tgtactgcag aatacatctg gtgtgtgat tttttttt aagcaagatt tgtttacta	4500
taataaagtg gattattca atgcaggcaa aattgtgaag tctgttggg aaagatagca	4560
tgtttttgt gtgaagtac ctgtcagta taagccttt tttttttt ttttaattta	4620
aatgtttgta gtgtctatgt ggacagttgt ttctagtgt ggtctgtgc ccaataactg	4680

[0010]

```

gggaacgagt tacagacaaa catcacgta aatgactcac aacattataa acagttgtga 4740
gaaaaatattt cacattatca aagctgtaca ataaa 4775

<210> 5
<211> 3162
<212> DNA
<213> 智人 (Homo sapiens)

<400> 5
atgtcagaggc gcaagcaggc gaaaccccag cacatcaact cggaggagga ccaggcgag 60
cagcagcgc agcagcagac cccgagitt gcagatggg cccagcggc gcccgcgcg 120
gggagctgg gtgctccagt gaaccacca gggaatgac aggtggcgag ttaggatgaa 180
gccacagtaa agcggtctcg tgggaggag acgcacgtct gtgagaaatg ctgtcggag 240
ttcttcagca tctctgagtt cctggaacat aagaaaaatt gcaactaaaa tccacctgc 300
ctcactatga atgacagcga ggggctgtg ccttcagaag acttctccgg agctgtactg 360
agccaccagc ccaccagtcc cggcagtaag gactgtcaca gggagaatgg cggcagctca 420
gaggacatga aggagaagcc ggaatgggag tctgtgggt acctaaagac agagacagcc 480
ctgccacca cccccagga cataagctat ttagecaaa gcaagtggc caacactaat 540
gtgaccttgc aggcactacg gggcaccaag gtggcggta atcagcggag cgcggatgca 600
ctccctccc ccgtgcctgg tgccaacagc atcccgtggg tcttcagca gatcttgtt 660
ctgcagcagc agcagctaca gcagatccag ctcacagagc agatccgcct ccagtgtaac 720
atgtgggect ccaacgect ccaactaagc ggggcagggg ccgacacict gaagacctg 780
ggcagccaca tgtctcagca gttttctgca gctgtggtt tgcacagcca gaaagctgga 840
agccaaggtc tctctctgga tgccttgaaa caagccaagc taactcagc caacatcct 900
tetgccacca gctccctgtc cccagggtg gcaccttca ctcgaagcc ggaagggaac 960
cgggtgtctc cgaacgtcat gtcgcctc ccgagcgtt tgccttctca ggcoccggc 1020
tcggtgtctt tccagagccc tttctccact gtggcgtag acacatccaa gaaagggaag 1080
gggaagccac cgaacatctc cgggtggat gtaaaccca aagacaggc ggcctctac 1140
aagcacaagt gtaagtaetg tagcaaggtt tttgggactg atagctcctt gcagatccac 1200
ctccgtctcc acactggaga gagaccttc gtgtgtctg tctgtgtgca tegtctcacc 1260
accaaggcca acctcaaggt gcactttcac cgacatccc aggtgaagcc aaaccccag 1320
ctgtttgcgc agttccagga caaagtggc gccggcaatg gcactcccta tgcactctt 1380
gtacctgacc ccatagatga accgagtctt tcttagaca gcaaacctgt ccttgtaacc 1440
acctctgtag ggtacctca gaatcttct tggggacta atcccaagga cctcaggggt 1500
ggtecttgc cgggtacct gcagctggg ccttctccag aaagtgggg tggaccaca 1560
ctccctgggg tgggaccaaa ctataattc ccaagggtg gtggcttcca agggagtggg 1620
accttgagc cagggtcaga gacctgaaa ttgcagcagt tggggagaa catgacaag 1680
gccaccacg aicccaaga atgtctcatt tgcacagag tctlaagctg tcagagctcc 1740
ctcaagatgc attatgcac ccacacggg gagagacct tccagtgtaa gatctgtgc 1800
cgagcctttt ctaccaagg taacctgaag acacacctg ggttccacc aaccaacaca 1860
tccattaga cgcagcttc gtgcccac tgcagaaga agtctactaa tgcctgatg 1920
ctgcagcaac atattcgat gcacatggc ggtcagatt ccaacagcc cctgcccag 1980
aatccctgtg actttacggg ttctgagcca atgacctgg gtgagaacg cagcaccggc 2040

```

[0011]

```

gctatctgcc atgatgatgt catcgaaagc atcgatgtag aggaagtcag ctcccaggag 2100
getcccagca gctctcccaa ggtcccacg cctcttccca gcatecactc ggcatcaccc 2160
acgetagggt ttgcatgat ggtctctta gatccccag ggaaagtggg tctgcccct 2220
ttaaaccctgc agcgccaggc cagcagagaa aacggctccg tggagagaga tggcttgacc 2280
aacgaatcat cctcgctgat gggagaccag gagtataga gccgaagccc agatatectg 2340
gaaaccacat ccttcaggc actctcccg gccaatagtc aagccgaaag catcaagtea 2400
aagtctcccg atgctgggag caaagcagag agctccgaga acagccgac tgagatggaa 2460
ggctcgagaa gtctccctc caegtctatc cgagccccc cgacctatgt caaggttgaa 2520
gttcttgga catctgtgg accctcgaca ttgtcccag ggaigacccc ttgttagca 2580
gccagccac gccgacaggc caagcaacat ggtgcacac gggtgggaa gaacttctg 2640
tctgttagcg ctcttcagat ccaagagcgg actcacactg gagagaagcc ttttgtgtg 2700
aacatttgtg ggcgagcttt taccacaaa ggcaacttaa aggttcacta catgacacac 2760
ggggcgaaac ataactcagc ccgcccgtga aggaagtggc ccacgagaa caccatggct 2820
ctgttaggta cggacggaaa aagagctca gaaatcttc ccaaggaaat cctggccct 2880
tcagtgaatg tggacctgt tgtgtggaac cagtacacca gcattctca tggcggtctg 2940
gccgtgaaga ccaatgagat ctctgtgac cagagtggg gggttctac cctccgggt 3000
tcttgggg ccacctccgt tgtgaataac gccactgtct ccaagatgga tggtcccag 3060
tcgggtatca gtgcagatgt gaaaaacca agtgctactg acggcgttc caaacaccag 3120
ttctctact tctggaaga aaacaagatt gcggtcagct aa 3162

```

<210> 6
 <211> 1851
 <212> DNA
 <213> 智人 (Homo sapiens)

<220>
 <221> 新的或少见的特征
 <222> (337)..(337)
 <223> n 是 a, c, g, 或 t

<220>
 <221> 新的或少见的特征
 <222> (421)..(421)
 <223> n 是 a, c, g, 或 t

<220>
 <221> 新的或少见的特征
 <222> (1317)..(1317)
 <223> n 是 a, c, g, 或 t

<220>
 <221> 新的或少见的特征
 <222> (1326)..(1326)
 <223> n 是 a, c, g, 或 t

<220>
 <221> 新的或少见的特征
 <222> (1331)..(1331)
 <223> n 是 a, c, g, 或 t

<220>
 <221> 新的或少见的特征
 <222> (1334)..(1334)
 <223> n 是 a, c, g, 或 t

<220>
 <221> 新的或少见的特征
 <222> (1367)..(1367)

[0012]

<223> n 是 a, c, g, 或 t

<220>

<221> 新的或少见的特征

<222> (1805).. (1805)

<223> n 是 a, c, g, 或 t

<220>

<221> 新的或少见的特征

<222> (1838).. (1838)

<223> n 是 a, c, g, 或 t

<400> 6

```

atgtctgaggc gcaagcaggc gaaaccccag cacatcaact oggaggagga ccagggcgag      60
cagcagccgc agcagcagac cccgagttt gcagatgcgg cccagcgcgc gcccgcgcg      120
ggggagctgg gtgtctcagt gaaccacca gggaaatgacg aggtggcgag tgaggatgaa      180
gccacagtaa agcggcttcg tgggaggag acgeacgtct gtgagaaatg ctgtgcggag      240
ttcttcagca tctctgagtt ctggaacat aagaaaaatt gcactaaaaa tccacctgic      300
ctcatcatga atgacagcga ggsgcctgtg ccttcanaag acttctccgg agctgtactg      360
agccaccagc ccaccagtc ccgcagttag gactgtcaca gggagaatgg ccgcagctca      420
naggacataa aggagaaagc ggaatgcggag tctgtgtgtt acctaaagac agagacagcc      480
ctgccaecca ccccccagga cataagctat ttageccaaag gcaaatggc caacactaac      540
glgacctlge agscactacg gggcaccag gggcgglga atcagcgag ccggaglcga      600
ctccctgccc ccgtgcctgg tcccacagc atcccgctgg tctctagca gatcttgtt      660
ctgcagcagc agcagctaca gcagatccag ctcccgagc agatccgat ccaggtgaac      720
atgtgggect cccacgccc cactcaagc ggggcaggg ccgacactct gaagacctg      780
ggcagccaca tgtctcagca ggtttctgca gctgtgctt tctctagca gaaagctgga      840
agccaagatc tgtctctgga tgccttgaaa caagccaaag tactcagc caacatccct      900
tctgccacca gctccctgtc cccagggtg gaaccttca ctctgaagc ggaatgggacc      960
cgggtgtctc cgaacgtcat gtcccgcct cccagcgctt tcttctctca ggccecggc      1020
tcggtgtctc tcagagccc ttctctcact gtggcgctag acacatccaa gaaagggaag      1080
gggaagccac cgaacatctc ccggttgat gtaaaccca aagacaggc ggccctctac      1140
aagcacaagt gtcggagcag tctctctcc acgtttatcc gagcccgcc gacctatgtc      1200
aaggttgaag ttcttggaac atttgtgga cctctgacat tctcccagg gatgacctc      1260
ttgttagcag cccagccag ccgacagcc aagcaacatg gctgcacag gtgtggnaag      1320
aacctnctgt ntgntagcgc tcttcagatc cagagcgga ctacantgg agagaagcct      1380
tttgtgtgca acatttgtg gcgagctttt accaccaag gcaactaaa ggttcactac      1440
atgacacacg gggcgaacaa taactcagc ccgcgtgga ggaagtggc catcgagaac      1500
accatggctc tgttaggtac ggaaggaaaa agagctcag aaatcttcc caaggaaac      1560
ctggccctt cagtgaatgt ggaacctgtt gtgtggaacc agtacaccag catgtcaat      1620
ggcggtctgg ccgtgaagac caatgagatc tctgtgctc agagtgggg ggttctctac      1680
ctcccggttt ccttggggc cactccgtt gtgaataacg caactgtct caagatggat      1740
ggctccagct cgggtatcag tgcagatgt gaaaaacca gtgtactga cggcgttccc      1800
aaacnceagt ttctcactt ctggaagaa aacaagantg cggtcagct a      1851

```

<210> 7

<211> 1320

<212> PRT

[0013]

<213> 智人(Homo sapiens)

<400> 7

Met Ser Arg Arg Lys Gln Ala Lys Pro Gln His Phe Gln Ser Asp Pro
1 5 10 15

Glu Val Ala Ser Leu Pro Arg Arg Asp Gly Asp Thr Glu Lys Gly Gln
20 25 30

Pro Ser Arg Pro Thr Lys Ser Lys Asp Ala His Val Cys Gly Arg Cys
35 40 45

Cys Ala Glu Phe Phe Glu Leu Ser Asp Leu Leu Leu His Lys Lys Asn
50 55 60

Cys Thr Lys Asn Gln Leu Val Leu Ile Val Asn Glu Asn Pro Ala Ser
65 70 75 80

Pro Pro Glu Thr Phe Ser Pro Ser Pro Pro Pro Asp Asn Pro Asp Glu
85 90 95

Gln Met Asn Asp Thr Val Asn Lys Thr Asp Gln Val Asp Cys Ser Asp
100 105 110

Leu Ser Glu His Asn Gly Leu Asp Arg Glu Glu Ser Met Glu Val Glu
115 120 125

Ala Pro Val Ala Asn Lys Ser Gly Ser Gly Thr Ser Ser Gly Ser His
130 135 140

Ser Ser Thr Ala Pro Ser Ser Ser Ser Ser Ser Ser Ser Ser Gly
145 150 155 160

Gly Gly Gly Ser Ser Ser Thr Gly Thr Ser Ala Ile Thr Thr Ser Leu
165 170 175

Pro Gln Leu Gly Asp Leu Thr Thr Leu Gly Asn Phe Ser Val Ile Asn
180 185 190

Ser Asn Val Ile Ile Glu Asn Leu Gln Ser Thr Lys Val Ala Val Ala
195 200 205

Gln Phe Ser Gln Glu Ala Arg Cys Gly Gly Ala Ser Gly Gly Lys Leu
210 215 220

Ala Val Pro Ala Leu Met Glu Gln Leu Leu Ala Leu Gln Gln Gln Gln
225 230 235 240

Ile His Gln Leu Gln Leu Ile Glu Gln Ile Arg His Gln Ile Leu Leu
245 250 255

Leu Ala Ser Gln Asn Ala Asp Leu Pro Thr Ser Ser Ser Pro Ser Gln
260 265 270

Gly Thr Leu Arg Thr Ser Ala Asn Pro Leu Ser Thr Leu Ser Ser His
275 280 285

Leu Ser Gln Gln Leu Ala Ala Ala Ala Gly Leu Ala Gln Ser Leu Ala
290 295 300

[0014]

Ser Gln Ser Ala Ser Ile Ser Gly Val Lys Gln Leu Pro Pro Ile Gln
 305 310 315 320
 Leu Pro Gln Ser Ser Ser Gly Asn Thr Ile Ile Pro Ser Asn Ser Gly
 325 330 335
 Ser Ser Pro Asn Met Asn Ile Leu Ala Ala Ala Val Thr Thr Pro Ser
 340 345 350
 Ser Glu Lys Val Ala Ser Ser Ala Gly Ala Ser His Val Ser Asn Pro
 355 360 365
 Ala Val Ser Ser Ser Ser Ser Pro Ala Phe Ala Ile Ser Ser Leu Leu
 370 375 380
 Ser Pro Ala Ser Asn Pro Leu Leu Pro Gln Gln Ala Ser Ala Asn Ser
 385 390 395 400
 Val Phe Pro Ser Pro Leu Pro Asn Ile Gly Thr Thr Ala Glu Asp Leu
 405 410 415
 Asn Ser Leu Ser Ala Leu Ala Gln Gln Arg Lys Ser Lys Pro Pro Asn
 420 425 430
 Val Thr Ala Phe Glu Ala Lys Ser Thr Ser Asp Glu Ala Phe Phe Lys
 435 440 445
 His Lys Cys Arg Phe Cys Ala Lys Val Phe Gly Ser Asp Ser Ala Leu
 450 455 460
 Gln Ile His Leu Arg Ser His Thr Gly Glu Arg Pro Phe Lys Cys Asn
 465 470 475 480
 Ile Cys Gly Asn Arg Phe Ser Thr Lys Gly Asn Leu Lys Val His Phe
 485 490 495
 Gln Arg His Lys Glu Lys Tyr Pro His Ile Gln Met Asn Pro Tyr Pro
 500 505 510
 Val Pro Glu His Leu Asp Asn Ile Pro Thr Ser Thr Gly Ile Pro Tyr
 515 520 525
 Gly Met Ser Ile Pro Pro Glu Lys Pro Val Thr Ser Trp Leu Asp Thr
 530 535 540
 Lys Pro Val Leu Pro Thr Leu Thr Thr Ser Val Gly Leu Pro Leu Pro
 545 550 555 560
 Pro Thr Leu Pro Ser Leu Ile Pro Phe Ile Lys Thr Glu Glu Pro Ala
 565 570 575
 Pro Ile Pro Ile Ser His Ser Ala Thr Ser Pro Pro Gly Ser Val Lys
 580 585 590
 Ser Asp Ser Gly Gly Pro Glu Ser Ala Thr Arg Asn Leu Gly Gly Leu
 595 600 605
 Pro Glu Glu Ala Glu Gly Ser Thr Leu Pro Pro Ser Gly Gly Lys Ser
 610 615 620

[0015]

Glu Glu Ser Gly Met Val Thr Asn Ser Val Pro Thr Ala Ser Ser Ser
 625 630 635 640
 Val Leu Ser Ser Pro Ala Ala Asp Cys Gly Pro Ala Gly Ser Ala Thr
 645 650 655
 Thr Phe Thr Asn Pro Leu Leu Pro Leu Met Ser Glu Gln Phe Lys Ala
 660 665 670
 Lys Phe Pro Phe Gly Gly Leu Leu Asp Ser Ala Gln Ala Ser Glu Thr
 675 680 685
 Ser Lys Leu Gln Gln Leu Val Glu Asn Ile Asp Lys Lys Ala Thr Asp
 690 695 700
 Pro Asn Glu Cys Ile Ile Cys His Arg Val Leu Ser Cys Gln Ser Ala
 705 710 715 720
 Leu Lys Met His Tyr Arg Thr His Thr Gly Glu Arg Pro Phe Lys Cys
 725 730 735
 Lys Ile Cys Gly Arg Ala Phe Thr Thr Lys Gly Asn Leu Lys Thr His
 740 745 750
 Tyr Ser Val His Arg Ala Met Pro Pro Leu Arg Val Gln His Ser Cys
 755 760 765
 Pro Ile Cys Gln Lys Lys Phe Thr Asn Ala Val Val Leu Gln Gln His
 770 775 780
 Ile Arg Met His Met Gly Gly Gln Ile Pro Asn Thr Pro Val Pro Asp
 785 790 795 800
 Ser Tyr Ser Glu Ser Met Glu Ser Asp Thr Gly Ser Phe Asp Glu Lys
 805 810 815
 Asn Phe Asp Asp Leu Asp Asn Phe Ser Asp Glu Asn Met Glu Asp Cys
 820 825 830
 Pro Glu Gly Ser Ile Pro Asp Thr Pro Lys Ser Ala Asp Ala Ser Gln
 835 840 845
 Asp Ser Leu Ser Ser Ser Pro Leu Pro Leu Glu Met Ser Ser Ile Ala
 850 855 860
 Ala Leu Glu Asn Gln Met Lys Met Ile Asn Ala Gly Leu Ala Glu Gln
 865 870 875 880
 Leu Gln Ala Ser Leu Lys Ser Val Glu Asn Gly Ser Ile Glu Gly Asp
 885 890 895
 Val Leu Thr Asn Asp Ser Ser Ser Val Gly Gly Asp Met Glu Ser Gln
 900 905 910
 Ser Ala Gly Ser Pro Ala Ile Ser Glu Ser Thr Ser Ser Met Gln Ala
 915 920 925
 Leu Ser Pro Ser Asn Ser Thr Gln Gln Phe His Lys Ser Pro Ser Ile

[0016]

930	935	940
Glu Glu Lys Pro Gln Arg Ala Val Pro Ser Glu Phe Ala Asn Gly Leu 945	950	955
Ser Pro Thr Pro Val Asn Gly Gly Ala Leu Asp Leu Thr Ser Ser His 965	970	975
Ala Glu Lys Ile Ile Lys Glu Asp Ser Leu Gly Ile Leu Phe Pro Phe 980	985	990
Arg Asp Arg Gly Lys Phe Lys Asn Thr Ala Cys Asp Ile Cys Gly Lys 995	1000	1005
Thr Phe Ala Cys Gln Ser Ala Leu Asp Ile His Tyr Arg Ser His 1010	1015	1020
Thr Lys Glu Arg Pro Phe Ile Cys Thr Val Cys Asn Arg Gly Phe 1025	1030	1035
Ser Thr Lys Gly Asn Leu Lys Gln His Met Leu Thr His Gln Met 1040	1045	1050
Arg Asp Leu Pro Ser Gln Leu Phe Glu Pro Ser Ser Asn Leu Gly 1055	1060	1065
Pro Asn Gln Asn Ser Ala Val Ile Pro Ala Asn Ser Leu Ser Ser 1070	1075	1080
Leu Ile Lys Thr Glu Val Asn Gly Phe Val His Val Ser Pro Gln 1085	1090	1095
Asp Ser Lys Asp Thr Pro Thr Ser His Val Pro Ser Gly Pro Leu 1100	1105	1110
Ser Ser Ser Ala Thr Ser Pro Val Leu Leu Pro Ala Leu Pro Arg 1115	1120	1125
Arg Thr Pro Lys Gln His Tyr Cys Asn Thr Cys Gly Lys Thr Phe 1130	1135	1140
Ser Ser Ser Ser Ala Leu Gln Ile His Glu Arg Thr His Thr Gly 1145	1150	1155
Glu Lys Pro Phe Ala Cys Thr Ile Cys Gly Arg Ala Phe Thr Thr 1160	1165	1170
Lys Gly Asn Leu Lys Val His Met Gly Thr His Met Trp Asn Ser 1175	1180	1185
Thr Pro Ala Arg Arg Gly Arg Arg Leu Ser Val Asp Gly Pro Met 1190	1195	1200
Thr Phe Leu Gly Gly Asn Pro Val Lys Phe Pro Glu Met Phe Gln 1205	1210	1215
Lys Asp Leu Ala Ala Arg Ser Gly Ser Gly Asp Pro Ser Ser Phe 1220	1225	1230

[0017]

Trp Asn Gln Tyr Ala Ala Ala Leu Ser Asn Gly Leu Ala Met Lys
 1235 1240 1245
 Ala Asn Glu Ile Ser Val Ile Gln Asn Gly Gly Ile Pro Pro Ile
 1250 1255 1260
 Pro Gly Ser Leu Gly Ser Gly Asn Ser Ser Pro Val Ser Gly Leu
 1265 1270 1275
 Thr Gly Asn Leu Glu Arg Leu Gln Asn Ser Glu Pro Asn Ala Pro
 1280 1285 1290
 Leu Ala Gly Leu Glu Lys Met Ala Ser Ser Glu Asn Gly Thr Asn
 1295 1300 1305
 Phe Arg Phe Thr Arg Phe Val Glu Asp Ser Lys Glu
 1310 1315 1320
 <210> 8
 <211> 1227
 <212> PRT
 <213> 智人(Homo sapiens)
 <400> 8
 Met Asn Asp Thr Val Asn Lys Thr Asp Gln Val Asp Cys Ser Asp Leu
 1 5 10 15
 Ser Glu His Asn Gly Leu Asp Arg Glu Glu Ser Met Glu Val Glu Ala
 20 25 30
 Pro Val Ala Asn Lys Ser Gly Ser Gly Thr Ser Ser Gly Ser His Ser
 35 40 45
 Ser Thr Ala Pro Ser Ser Ser Ser Ser Ser Ser Ser Ser Gly Gly
 50 55 60
 Gly Gly Ser Ser Ser Thr Gly Thr Ser Ala Ile Thr Thr Ser Leu Pro
 65 70 75 80
 Gln Leu Gly Asp Leu Thr Thr Leu Gly Asn Phe Ser Val Ile Asn Ser
 85 90 95
 Asn Val Ile Ile Glu Asn Leu Gln Ser Thr Lys Val Ala Val Ala Gln
 100 105 110
 Phe Ser Gln Glu Ala Arg Cys Gly Gly Ala Ser Gly Gly Lys Leu Ala
 115 120 125
 Val Pro Ala Leu Met Glu Gln Leu Leu Ala Leu Gln Gln Gln Ile
 130 135 140
 His Gln Leu Gln Leu Ile Gln Gln Ile Arg His Gln Ile Leu Leu Leu
 145 150 155 160
 Ala Ser Gln Asn Ala Asp Leu Pro Thr Ser Ser Ser Pro Ser Gln Gly
 165 170 175
 Thr Leu Arg Thr Ser Ala Asn Pro Leu Ser Thr Leu Ser Ser His Leu
 180 185 190

[0018]

```

Ser Gln Gln Leu Ala Ala Ala Gly Leu Ala Gln Ser Leu Ala Ser
195                200                205

Gln Ser Ala Ser Ile Ser Gly Val Lys Gln Leu Pro Pro Ile Gln Leu
210                215                220

Pro Gln Ser Ser Ser Gly Asn Thr Ile Ile Pro Ser Asn Ser Gly Ser
225                230                235                240

Ser Pro Asn Met Asn Ile Leu Ala Ala Ala Val Thr Thr Pro Ser Ser
245                250                255

Glu Lys Val Ala Ser Ser Ala Gly Ala Ser His Val Ser Asn Pro Ala
260                265                270

Val Ser Ser Ser Ser Ser Pro Ala Phe Ala Ile Ser Ser Leu Leu Ser
275                280                285

Pro Ala Ser Asn Pro Leu Leu Pro Gln Gln Ala Ser Ala Asn Ser Val
290                295                300

Phe Pro Ser Pro Leu Pro Asn Ile Gly Thr Thr Ala Glu Asp Leu Asn
305                310                315                320

Ser Leu Ser Ala Leu Ala Gln Gln Arg Lys Ser Lys Pro Pro Asn Val
325                330                335

Thr Ala Phe Glu Ala Lys Ser Thr Ser Asp Glu Ala Phe Phe Lys His
340                345                350

Lys Cys Arg Phe Cys Ala Lys Val Phe Gly Ser Asp Ser Ala Leu Gln
355                360                365

Ile His Leu Arg Ser His Thr Gly Glu Arg Pro Phe Lys Cys Asn Ile
370                375                380

Cys Gly Asn Arg Phe Ser Thr Lys Gly Asn Leu Lys Val His Phe Gln
385                390                395                400

Arg His Lys Glu Lys Tyr Pro His Ile Gln Met Asn Pro Tyr Pro Val
405                410                415

Pro Gln His Leu Asp Asn Ile Pro Thr Ser Thr Gly Ile Pro Tyr Gly
420                425                430

Met Ser Ile Pro Pro Glu Lys Pro Val Thr Ser Trp Leu Asp Thr Lys
435                440                445

Pro Val Leu Pro Thr Leu Thr Thr Ser Val Gly Leu Pro Leu Pro Pro
450                455                460

Thr Leu Pro Ser Leu Ile Pro Phe Ile Lys Thr Glu Glu Pro Ala Pro
465                470                475                480

Ile Pro Ile Ser His Ser Ala Thr Ser Pro Pro Gly Ser Val Lys Ser
485                490                495

Asp Ser Gly Gly Pro Glu Ser Ala Thr Arg Asn Leu Gly Gly Leu Pro
500                505                510

```

[0019]

Glu Glu Ala Glu Gly Ser Thr Leu Pro Pro Ser Gly Gly Lys Ser Glu
 515 520 525
 Glu Ser Gly Met Val Thr Asn Ser Val Pro Thr Ala Ser Ser Ser Val
 530 535 540
 Leu Ser Ser Pro Ala Ala Asp Cys Gly Pro Ala Gly Ser Ala Thr Thr
 545 550 555 560
 Phe Thr Asn Pro Leu Leu Pro Leu Met Ser Glu Gln Phe Lys Ala Lys
 565 570 575
 Phe Pro Phe Gly Gly Leu Leu Asp Ser Ala Gln Ala Ser Glu Thr Ser
 580 585 590
 Lys Leu Glu Gln Leu Val Glu Asn Ile Asp Lys Lys Ala Thr Asp Pro
 595 600 605
 Asn Glu Cys Ile Ile Cys His Arg Val Leu Ser Cys Gln Ser Ala Leu
 610 615 620
 Lys Met His Tyr Arg Thr His Thr Gly Glu Arg Pro Phe Lys Cys Lys
 625 630 635 640
 Ile Cys Gly Arg Ala Phe Thr Thr Lys Gly Asn Leu Lys Thr His Tyr
 645 650 655
 Ser Val His Arg Ala Met Pro Pro Leu Arg Val Gln His Ser Cys Pro
 660 665 670
 Ile Cys Gln Lys Lys Phe Thr Asn Ala Val Val Leu Gln Gln His Ile
 675 680 685
 Arg Met His Met Gly Gly Gln Ile Pro Asn Thr Pro Val Pro Asp Ser
 690 695 700
 Tyr Ser Glu Ser Met Glu Ser Asp Thr Gly Ser Phe Asp Glu Lys Asn
 705 710 715 720
 Phe Asp Asp Leu Asp Asn Phe Ser Asp Glu Asn Met Glu Asp Cys Pro
 725 730 735
 Glu Gly Ser Ile Pro Asp Thr Pro Lys Ser Ala Asp Ala Ser Gln Asp
 740 745 750
 Ser Leu Ser Ser Ser Pro Leu Pro Leu Glu Met Ser Ser Ile Ala Ala
 755 760 765
 Leu Glu Asn Gln Met Lys Met Ile Asn Ala Gly Leu Ala Glu Gln Leu
 770 775 780
 Gln Ala Ser Leu Lys Ser Val Glu Asn Gly Ser Ile Glu Gly Asp Val
 785 790 795 800
 Leu Thr Asn Asp Ser Ser Ser Val Gly Gly Asp Met Glu Ser Gln Ser
 805 810 815
 Ala Gly Ser Pro Ala Ile Ser Glu Ser Thr Ser Ser Met Gln Ala Leu
 820 825 830

[0020]

Ser Pro Ser Asn Ser Thr Gln Glu Phe His Lys Ser Pro Ser Ile Glu
 835 840 845
 Glu Lys Pro Gln Arg Ala Val Pro Ser Glu Phe Ala Asn Gly Leu Ser
 850 855 860
 Pro Thr Pro Val Asn Gly Gly Ala Leu Asp Leu Thr Ser Ser His Ala
 865 870 875 880
 Glu Lys Ile Ile Lys Glu Asp Ser Leu Gly Ile Leu Phe Pro Phe Arg
 885 890 895
 Asp Arg Gly Lys Phe Lys Asn Thr Ala Cys Asp Ile Cys Gly Lys Thr
 900 905 910
 Phe Ala Cys Gln Ser Ala Leu Asp Ile His Tyr Arg Ser His Thr Lys
 915 920 925
 Glu Arg Pro Phe Ile Cys Thr Val Cys Asn Arg Gly Phe Ser Thr Lys
 930 935 940
 Gly Asn Leu Lys Gln His Met Leu Thr His Gln Met Arg Asp Leu Pro
 945 950 955 960
 Ser Gln Leu Phe Glu Pro Ser Ser Asn Leu Gly Pro Asn Gln Asn Ser
 965 970 975
 Ala Val Ile Pro Ala Asn Ser Leu Ser Ser Leu Ile Lys Thr Glu Val
 980 985 990
 Asn Gly Phe Val His Val Ser Pro Gln Asp Ser Lys Asp Thr Pro Thr
 995 1000 1005
 Ser His Val Pro Ser Gly Pro Leu Ser Ser Ser Ala Thr Ser Pro
 1010 1015 1020
 Val Leu Leu Pro Ala Leu Pro Arg Arg Thr Pro Lys Gln His Tyr
 1025 1030 1035
 Cys Asn Thr Cys Gly Lys Thr Phe Ser Ser Ser Ser Ala Leu Gln
 1040 1045 1050
 Ile His Glu Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Thr
 1055 1060 1065
 Ile Cys Gly Arg Ala Phe Thr Thr Lys Gly Asn Leu Lys Val His
 1070 1075 1080
 Met Gly Thr His Met Trp Asn Ser Thr Pro Ala Arg Arg Gly Arg
 1085 1090 1095
 Arg Leu Ser Val Asp Gly Pro Met Thr Phe Leu Gly Gly Asn Pro
 1100 1105 1110
 Val Lys Phe Pro Glu Met Phe Gln Lys Asp Leu Ala Ala Arg Ser
 1115 1120 1125
 Gly Ser Gly Asp Pro Ser Ser Phe Trp Asn Gln Tyr Ala Ala Ala

[0021]

1130	1135	1140
Leu Ser Asn Gly Leu Ala Met	Lys Ala Asn Glu Ile	Ser Val Ile
1145	1150	1155
Gln Asn Gly Gly Ile Pro Pro	Ile Pro Gly Ser Leu	Gly Ser Gly
1160	1165	1170
Asn Ser Ser Pro Val Ser Gly	Leu Thr Gly Asn Leu	Glu Arg Leu
1175	1180	1185
Gln Asn Ser Glu Pro Asn Ala	Pro Leu Ala Gly Leu	Glu Lys Met
1190	1195	1200
Ala Ser Ser Glu Asn Gly Thr	Asn Phe Arg Phe Thr	Arg Phe Val
1205	1210	1215
Glu Asp Ser Lys Glu Ile Val	Thr Ser	
1220	1225	
<210> 9		
<211> 1007		
<212> PRT		
<213> 智人(Homo sapiens)		
<400> 9		
Met Ser Arg Arg Lys Gln Arg Lys Pro Gln Gln Leu Ile Ser Asp Cys		
1 5 10 15		
Glu Gly Pro Ser Ala Ser Glu Asn Gly Asp Ala Ser Glu Glu Asp His		
20 25 30		
Pro Gln Val Cys Ala Lys Cys Cys Ala Gln Phe Thr Asp Pro Thr Glu		
35 40 45		
Phe Leu Ala His Gln Asn Ala Cys Ser Thr Asp Pro Pro Val Met Val		
50 55 60		
Ile Ile Gly Gly Gln Glu Asn Pro Asn Asn Ser Ser Ala Ser Ser Glu		
65 70 75 80		
Pro Arg Pro Glu Gly His Asn Asn Pro Gln Val Met Asp Thr Glu His		
85 90 95		
Ser Asn Pro Pro Asp Ser Gly Ser Ser Val Pro Thr Asp Pro Thr Trp		
100 105 110		
Gly Pro Glu Arg Arg Gly Glu Glu Ser Pro Gly His Phe Leu Val Ala		
115 120 125		
Ala Thr Gly Thr Ala Ala Gly Gly Gly Gly Gly Leu Ile Leu Ala Ser		
130 135 140		
Pro Lys Leu Gly Ala Thr Pro Leu Pro Pro Glu Ser Thr Pro Ala Pro		
145 150 155 160		
Pro Pro Pro Pro Pro Pro Pro Pro Gly Val Gly Ser Gly His		
165 170 175		
Leu Asn Ile Pro Leu Ile Leu Glu Glu Leu Arg Val Leu Gln Gln Arg		

[0022]

180	185	190
Gln Ile His Gln Met Gln Met Thr Glu Gln Ile Cys Arg Gln Val Leu 195	200	205
Leu Leu Gly Ser Leu Gly Gln Thr Val Gly Ala Pro Ala Ser Pro Ser 210	215	220
Glu Leu Pro Gly Thr Gly Thr Ala Ser Ser Thr Lys Pro Leu Leu Pro 225	230	235
Leu Phe Ser Pro Ile Lys Pro Val Gln Thr Ser Lys Thr Leu Ala Ser 245	250	255
Ser Ser Ser Ser Ser Ser Ser Ser Ser Gly Ala Glu Thr Pro Lys Gln 260	265	270
Ala Phe Phe His Leu Tyr His Pro Leu Gly Ser Gln His Pro Phe Ser 275	280	285
Ala Gly Gly Val Gly Arg Ser His Lys Pro Thr Pro Ala Pro Ser Pro 290	295	300
Ala Leu Pro Gly Ser Thr Asp Gln Leu Ile Ala Ser Pro His Leu Ala 305	310	315
Phe Pro Ser Thr Thr Gly Leu Leu Ala Ala Gln Cys Leu Gly Ala Ala 325	330	335
Arg Gly Leu Glu Ala Thr Ala Ser Pro Gly Leu Leu Lys Pro Lys Asn 340	345	350
Gly Ser Gly Glu Leu Ser Tyr Gly Glu Val Met Gly Pro Leu Glu Lys 355	360	365
Pro Gly Gly Arg His Lys Cys Arg Phe Cys Ala Lys Val Phe Gly Ser 370	375	380
Asp Ser Ala Leu Gln Ile His Leu Arg Ser His Thr Gly Glu Arg Pro 385	390	395
Tyr Lys Cys Asn Val Cys Gly Asn Arg Phe Thr Thr Arg Gly Asn Leu 405	410	415
Lys Val His Phe His Arg His Arg Glu Lys Tyr Pro His Val Gln Met 420	425	430
Asn Pro His Pro Val Pro Glu His Leu Asp Tyr Val Ile Thr Ser Ser 435	440	445
Gly Leu Pro Tyr Gly Met Ser Val Pro Pro Glu Lys Ala Glu Glu Glu 450	455	460
Ala Ala Thr Pro Gly Gly Gly Val Glu Arg Lys Pro Leu Val Ala Ser 465	470	475
Thr Thr Ala Leu Ser Ala Thr Glu Ser Leu Thr Leu Leu Ser Thr Ser 485	490	495

[0023]

Ala Gly Thr Ala Thr Ala Pro Gly Leu Pro Ala Phe Asn Lys Phe Val
 500 505 510
 Leu Met Lys Ala Val Glu Pro Lys Asn Lys Ala Asp Glu Asn Thr Pro
 515 520 525
 Pro Gly Ser Glu Gly Ser Ala Ile Ser Gly Val Ala Glu Ser Ser Thr
 530 535 540
 Ala Thr Arg Met Gln Leu Ser Lys Leu Val Thr Ser Leu Pro Ser Trp
 545 550 555 560
 Ala Leu Leu Thr Asn His Phe Lys Ser Thr Gly Ser Phe Pro Phe Pro
 565 570 575
 Tyr Val Leu Glu Pro Leu Gly Ala Ser Pro Ser Glu Thr Ser Lys Leu
 580 585 590
 Gln Gln Leu Val Glu Lys Ile Asp Arg Gln Gly Ala Val Ala Val Thr
 595 600 605
 Ser Ala Ala Ser Gly Ala Pro Thr Thr Ser Ala Pro Ala Pro Ser Ser
 610 615 620
 Ser Ala Ser Ser Gly Pro Asn Gln Cys Val Ile Cys Leu Arg Val Leu
 625 630 635 640
 Ser Cys Pro Arg Ala Leu Arg Leu His Tyr Gly Gln His Gly Gly Glu
 645 650 655
 Arg Pro Phe Lys Cys Lys Val Cys Gly Arg Ala Phe Ser Thr Arg Gly
 660 665 670
 Asn Leu Arg Ala His Phe Val Gly His Lys Ala Ser Pro Ala Ala Arg
 675 680 685
 Ala Gln Asn Ser Cys Pro Ile Cys Gln Lys Lys Phe Thr Asn Ala Val
 690 695 700
 Thr Leu Gln Gln His Val Arg Met His Leu Gly Gly Gln Ile Pro Asn
 705 710 715 720
 Gly Gly Thr Ala Leu Pro Glu Gly Gly Gly Ala Ala Gln Glu Asn Gly
 725 730 735
 Ser Glu Gln Ser Thr Val Ser Gly Ala Gly Ser Phe Pro Gln Gln Gln
 740 745 750
 Ser Gln Gln Pro Ser Pro Glu Glu Glu Leu Ser Glu Glu Glu Glu
 755 760 765
 Glu Asp Glu Glu Glu Glu Asp Val Thr Asp Glu Asp Ser Leu Ala
 770 775 780
 Gly Arg Gly Ser Glu Ser Gly Gly Glu Lys Ala Ile Ser Val Arg Gly
 785 790 795 800
 Asp Ser Glu Glu Ala Ser Gly Ala Glu Glu Glu Val Gly Thr Val Ala
 805 810 815

[0024]

Ala Ala Ala Thr Ala Gly Lys Glu Met Asp Ser Asn Glu Lys Thr Thr
 820 825 830
 Gln Gln Ser Ser Leu Pro Pro Pro Pro Pro Pro Asp Ser Leu Asp Gln
 835 840 845
 Pro Gln Pro Met Glu Gln Gly Ser Ser Gly Val Leu Gly Gly Lys Glu
 850 855 860
 Glu Gly Gly Lys Pro Glu Arg Ser Ser Ser Pro Ala Ser Ala Leu Thr
 865 870 875 880
 Pro Glu Gly Glu Ala Thr Ser Val Thr Leu Val Glu Glu Leu Ser Leu
 885 890 895
 Gln Glu Ala Met Arg Lys Glu Pro Gly Glu Ser Ser Ser Arg Lys Ala
 900 905 910
 Cys Glu Val Cys Gly Gln Ala Phe Pro Ser Gln Ala Ala Leu Glu Glu
 915 920 925
 His Gln Lys Thr His Pro Lys Glu Gly Pro Leu Phe Thr Cys Val Phe
 930 935 940
 Cys Arg Gln Gly Phe Leu Glu Arg Ala Thr Leu Lys Lys His Met Leu
 945 950 955 960
 Leu Ala His His Gln Val Gln Pro Phe Ala Pro His Gly Pro Gln Asn
 965 970 975
 Ile Ala Ala Leu Ser Leu Val Pro Gly Cys Ser Pro Ser Ile Thr Ser
 980 985 990
 Thr Gly Leu Ser Pro Phe Pro Arg Lys Asp Asp Pro Thr Ile Pro
 995 1000 1005
 <210> 10
 <211> 1300
 <212> PRT
 <213> 智人(Homo sapiens)
 <400> 10
 Met Ser Arg Arg Lys Gln Ala Lys Pro Gln His Leu Lys Ser Asp Glu
 1 5 10 15
 Glu Leu Leu Pro Pro Asp Gly Ala Pro Glu His Ala Ala Pro Gly Glu
 20 25 30
 Gly Ala Glu Asp Ala Asp Ser Gly Pro Glu Ser Arg Ser Gly Gly Glu
 35 40 45
 Glu Thr Ser Val Cys Glu Lys Cys Cys Ala Glu Phe Phe Lys Trp Ala
 50 55 60
 Asp Phe Leu Glu His Gln Arg Ser Cys Thr Lys Leu Pro Pro Val Leu
 65 70 75 80
 Ile Val His Glu Asp Ala Pro Ala Pro Pro Pro Glu Asp Phe Pro Glu
 85 90 95

[0025]

```

Pro Ser Pro Ala Ser Ser Pro Ser Glu Arg Ala Glu Ser Glu Ala Ala
100 105 110

Glu Glu Ala Gly Ala Glu Gly Ala Glu Gly Glu Ala Arg Pro Val Glu
115 120 125

Lys Glu Ala Glu Pro Met Asp Ala Glu Pro Ala Gly Asp Thr Arg Ala
130 135 140

Pro Arg Pro Pro Pro Ala Ala Pro Ala Pro Thr Pro Ala Tyr Gly
145 150 155 160

Ala Pro Ser Thr Asn Val Thr Leu Glu Ala Leu Leu Ser Thr Lys Val
165 170 175

Ala Val Ala Gln Phe Ser Gln Gly Ala Arg Ala Ala Gly Gly Ser Gly
180 185 190

Ala Gly Gly Gly Val Ala Ala Ala Ala Val Pro Leu Ile Leu Glu Gln
195 200 205

Leu Met Ala Leu Gln Gln Gln Gln Ile His Gln Leu Gln Leu Ile Glu
210 215 220

Gln Ile Arg Ser Gln Val Ala Leu Met Gln Arg Pro Pro Pro Arg Pro
225 230 235 240

Ser Leu Ser Pro Ala Ala Ala Pro Ser Ala Pro Gly Pro Ala Pro Ser
245 250 255

Gln Leu Pro Gly Leu Ala Ala Leu Pro Leu Ser Ala Gly Ala Pro Ala
260 265 270

Ala Ala Ile Ala Gly Ser Gly Pro Ala Ala Pro Ala Ala Phe Glu Gly
275 280 285

Ala Gln Pro Leu Ser Arg Pro Glu Ser Gly Ala Ser Thr Pro Gly Gly
290 295 300

Pro Ala Glu Pro Ser Ala Pro Ala Ala Pro Ser Ala Ala Pro Ala Pro
305 310 315 320

Ala Ala Pro Ala Pro Ala Pro Ala Pro Gln Ser Ala Ala Ser Ser Gln
325 330 335

Pro Gln Ser Ala Ser Thr Pro Pro Ala Leu Ala Pro Gly Ser Leu Leu
340 345 350

Gly Ala Ala Pro Gly Leu Pro Ser Pro Leu Leu Pro Gln Thr Ser Ala
355 360 365

Ser Gly Val Ile Phe Pro Asn Pro Leu Val Ser Ile Ala Ala Thr Ala
370 375 380

Asn Ala Leu Asp Pro Leu Ser Ala Leu Met Lys His Arg Lys Gly Lys
385 390 395 400

Pro Pro Asn Val Ser Val Phe Glu Pro Lys Ala Ser Ala Glu Asp Pro
405 410 415

```

[0026]

Phe Phe Lys His Lys Cys Arg Phe Cys Ala Lys Val Phe Gly Ser Asp
 420 425 430
 Ser Ala Leu Gln Ile His Leu Arg Ser His Thr Gly Glu Arg Pro Phe
 435 440 445
 Lys Cys Asn Ile Cys Gly Asn Arg Phe Ser Thr Lys Gly Asn Leu Lys
 450 455 460
 Val His Phe Gln Arg His Lys Glu Lys Tyr Pro His Ile Gln Met Asn
 465 470 475 480
 Pro Tyr Pro Val Pro Glu Tyr Leu Asp Asn Val Pro Thr Cys Ser Gly
 485 490 495
 Ile Pro Tyr Gly Met Ser Leu Pro Pro Glu Lys Pro Val Thr Thr Trp
 500 505 510
 Leu Asp Ser Lys Pro Val Leu Pro Thr Val Pro Thr Ser Val Gly Leu
 515 520 525
 Gln Leu Pro Pro Thr Val Pro Gly Ala His Gly Tyr Ala Asp Ser Pro
 530 535 540
 Ser Ala Thr Pro Ala Ser Arg Ser Pro Gln Arg Pro Ser Pro Ala Ser
 545 550 555 560
 Ser Glu Cys Ala Ser Leu Ser Pro Gly Leu Asn His Val Glu Ser Gly
 565 570 575
 Val Ser Ala Thr Ala Glu Ser Pro Gln Ser Leu Leu Gly Gly Pro Pro
 580 585 590
 Leu Thr Lys Ala Glu Pro Val Ser Leu Pro Cys Thr Asn Ala Arg Ala
 595 600 605
 Gly Asp Ala Pro Val Gly Ala Gln Ala Ser Ala Ala Pro Thr Ser Val
 610 615 620
 Asp Gly Ala Pro Thr Ser Leu Gly Ser Pro Gly Leu Pro Ala Val Ser
 625 630 635 640
 Glu Gln Phe Lys Ala Gln Phe Pro Phe Gly Gly Leu Leu Asp Ser Met
 645 650 655
 Gln Thr Ser Glu Thr Ser Lys Leu Gln Gln Leu Val Glu Asn Ile Asp
 660 665 670
 Lys Lys Met Thr Asp Pro Asn Gln Cys Val Ile Cys His Arg Val Leu
 675 680 685
 Ser Cys Gln Ser Ala Leu Lys Met His Tyr Arg Thr His Thr Gly Glu
 690 695 700
 Arg Pro Phe Lys Cys Lys Ile Cys Gly Arg Ala Phe Thr Thr Lys Gly
 705 710 715 720
 Asn Leu Lys Thr His Phe Gly Val His Arg Ala Lys Pro Pro Leu Arg

[0027]

725	730	735
Val Gln His Ser Cys Pro Ile Cys Gln Lys Lys Phe Thr Asn Ala Val 740	745	750
Val Leu Gln Gln His Ile Arg Met His Met Gly Gly Gln Ile Pro Asn 755	760	765
Thr Pro Leu Pro Glu Gly Phe Gln Asp Ala Met Asp Ser Glu Leu Ala 770	775	780
Tyr Asp Asp Lys Asn Ala Glu Thr Leu Ser Ser Tyr Asp Asp Asp Met 785	790	795
Asp Glu Asn Ser Met Glu Asp Asp Ala Glu Leu Lys Asp Ala Ala Thr 805	810	815
Asp Pro Ala Lys Pro Leu Leu Ser Tyr Ala Gly Ser Cys Pro Pro Ser 820	825	830
Pro Pro Ser Val Ile Ser Ser Ile Ala Ala Leu Glu Asn Gln Met Lys 835	840	845
Met Ile Asp Ser Val Met Ser Cys Gln Gln Leu Thr Gly Leu Lys Ser 850	855	860
Val Glu Asn Gly Ser Gly Glu Ser Asp Arg Leu Ser Asn Asp Ser Ser 865	870	875
Ser Ala Val Gly Asp Leu Glu Ser Arg Ser Ala Gly Ser Pro Ala Leu 885	890	895
Ser Glu Ser Ser Ser Ser Gln Ala Leu Ser Pro Ala Pro Ser Asn Gly 900	905	910
Glu Ser Phe Arg Ser Lys Ser Pro Gly Leu Gly Ala Pro Glu Glu Pro 915	920	925
Gln Glu Ile Pro Leu Lys Thr Glu Arg Pro Asp Ser Pro Ala Ala Ala 930	935	940
Pro Gly Ser Gly Gly Ala Pro Gly Arg Ala Gly Ile Lys Glu Glu Ala 945	950	955
Pro Phe Ser Leu Leu Phe Leu Ser Arg Glu Arg Gly Lys Cys Pro Ser 965	970	975
Thr Val Cys Gly Val Cys Gly Lys Pro Phe Ala Cys Lys Ser Ala Leu 980	985	990
Glu Ile His Tyr Arg Ser His Thr Lys Glu Arg Pro Phe Val Cys Ala 995	1000	1005
Leu Cys Arg Arg Gly Cys Ser Thr Met Gly Asn Leu Lys Gln His 1010	1015	1020
Leu Leu Thr His Arg Leu Lys Glu Leu Pro Ser Gln Leu Phe Asp 1025	1030	1035

[0028]

Pro Asn Phe Ala Leu Gly Pro Ser Gln Ser Thr Pro Ser Leu Ile
 1040 1045 1050
 Ser Ser Ala Ala Pro Thr Met Ile Lys Met Glu Val Asn Gly His
 1055 1060 1065
 Gly Lys Ala Met Ala Leu Gly Glu Gly Pro Pro Leu Pro Ala Gly
 1070 1075 1080
 Val Gln Val Pro Ala Gly Pro Gln Thr Val Met Gly Pro Gly Leu
 1085 1090 1095
 Ala Pro Met Leu Ala Pro Pro Pro Arg Arg Thr Pro Lys Gln His
 1100 1105 1110
 Asn Cys Gln Ser Cys Gly Lys Thr Phe Ser Ser Ala Ser Ala Leu
 1115 1120 1125
 Gln Ile His Glu Arg Thr His Thr Gly Glu Lys Pro Phe Gly Cys
 1130 1135 1140
 Thr Ile Cys Gly Arg Ala Phe Thr Thr Lys Gly Asn Leu Lys Val
 1145 1150 1155
 His Met Gly Thr His Met Trp Asn Asn Ala Pro Ala Arg Arg Gly
 1160 1165 1170
 Arg Arg Leu Ser Val Glu Asn Pro Met Ala Leu Leu Gly Gly Asp
 1175 1180 1185
 Ala Leu Lys Phe Ser Glu Met Phe Gln Lys Asp Leu Ala Ala Arg
 1190 1195 1200
 Ala Met Asn Val Asp Pro Ser Phe Trp Asn Gln Tyr Ala Ala Ala
 1205 1210 1215
 Ile Thr Asn Gly Leu Ala Met Lys Asn Asn Glu Ile Ser Val Ile
 1220 1225 1230
 Gln Asn Gly Gly Ile Pro Gln Leu Pro Val Ser Leu Gly Gly Ser
 1235 1240 1245
 Ala Leu Pro Pro Leu Gly Ser Met Ala Ser Gly Met Asp Lys Ala
 1250 1255 1260
 Arg Thr Gly Ser Ser Pro Pro Ile Val Ser Leu Asp Lys Ala Ser
 1265 1270 1275
 Ser Glu Thr Ala Ala Ser Arg Pro Phe Thr Arg Phe Ile Glu Asp
 1280 1285 1290
 Asn Lys Glu Ile Gly Ile Asn
 1295 1300
 <210> II
 <211> 1053
 <212> PRT
 <213> 智人(Homo sapiens)
 <400> II

[0029]

```

Met Ser Arg Arg Lys Gln Ala Lys Pro Gln His Ile Asn Ser Glu Glu
1      5      10      15

Asp Gln Gly Glu Gln Gln Pro Gln Gln Gln Thr Pro Glu Phe Ala Asp
      20      25      30

Ala Ala Pro Ala Ala Pro Ala Ala Gly Glu Leu Gly Ala Pro Val Asn
      35      40      45

His Pro Gly Asn Asp Glu Val Ala Ser Glu Asp Glu Ala Thr Val Lys
      50      55      60

Arg Leu Arg Arg Glu Glu Thr His Val Cys Glu Lys Cys Cys Ala Glu
      65      70      75      80

Phe Phe Ser Ile Ser Glu Phe Leu Glu His Lys Lys Asn Cys Thr Lys
      85      90      95

Asn Pro Pro Val Leu Ile Met Asn Asp Ser Glu Gly Pro Val Pro Ser
      100      105      110

Glu Asp Phe Ser Gly Ala Val Leu Ser His Gln Pro Thr Ser Pro Gly
      115      120      125

Ser Lys Asp Cys His Arg Glu Asn Gly Gly Ser Ser Glu Asp Met Lys
      130      135      140

Glu Lys Pro Asp Ala Glu Ser Val Val Tyr Leu Lys Thr Glu Thr Ala
      145      150      155      160

Leu Pro Pro Thr Pro Gln Asp Ile Ser Tyr Leu Ala Lys Gly Lys Val
      165      170      175

Ala Asn Thr Asn Val Thr Leu Gln Ala Leu Arg Gly Thr Lys Val Ala
      180      185      190

Val Asn Gln Arg Ser Ala Asp Ala Leu Pro Ala Pro Val Pro Gly Ala
      195      200      205

Asn Ser Ile Pro Trp Val Leu Glu Gln Ile Leu Cys Leu Gln Gln Gln
      210      215      220

Gln Leu Gln Gln Ile Gln Leu Thr Glu Gln Ile Arg Ile Gln Val Asn
      225      230      235      240

Met Trp Ala Ser His Ala Leu His Ser Ser Gly Ala Gly Ala Asp Thr
      245      250      255

Leu Lys Thr Leu Gly Ser His Met Ser Gln Gln Val Ser Ala Ala Val
      260      265      270

Ala Leu Leu Ser Gln Lys Ala Gly Ser Gln Gly Leu Ser Leu Asp Ala
      275      280      285

Leu Lys Gln Ala Lys Leu Pro His Ala Asn Ile Pro Ser Ala Thr Ser
      290      295      300

Ser Leu Ser Pro Gly Leu Ala Pro Phe Thr Leu Lys Pro Asp Gly Thr
      305      310      315      320

```

[0030]

Arg Val Leu Pro Asn Val Met Ser Arg Leu Pro Ser Ala Leu Leu Pro
 325 330 335
 Gln Ala Pro Gly Ser Val Leu Phe Gln Ser Pro Phe Ser Thr Val Ala
 340 345 350
 Leu Asp Thr Ser Lys Lys Gly Lys Gly Lys Pro Pro Asn Ile Ser Ala
 355 360 365
 Val Asp Val Lys Pro Lys Asp Glu Ala Ala Leu Tyr Lys His Lys Cys
 370 375 380
 Lys Tyr Cys Ser Lys Val Phe Gly Thr Asp Ser Ser Leu Gln Ile His
 385 390 395 400
 Leu Arg Ser His Thr Gly Glu Arg Pro Phe Val Cys Ser Val Cys Gly
 405 410 415
 His Arg Phe Thr Thr Lys Gly Asn Leu Lys Val His Phe His Arg His
 420 425 430
 Pro Gln Val Lys Ala Asn Pro Gln Leu Phe Ala Glu Phe Gln Asp Lys
 435 440 445
 Val Ala Ala Gly Asn Gly Ile Pro Tyr Ala Leu Ser Val Pro Asp Pro
 450 455 460
 Ile Asp Glu Pro Ser Leu Ser Leu Asp Ser Lys Pro Val Leu Val Thr
 465 470 475 480
 Thr Ser Val Gly Leu Pro Gln Asn Leu Ser Ser Gly Thr Asn Pro Lys
 485 490 495
 Asp Leu Thr Gly Gly Ser Leu Pro Gly Asp Leu Gln Pro Gly Pro Ser
 500 505 510
 Pro Glu Ser Glu Gly Gly Pro Thr Leu Pro Gly Val Gly Pro Asn Tyr
 515 520 525
 Asn Ser Pro Arg Ala Gly Gly Phe Gln Gly Ser Gly Thr Pro Glu Pro
 530 535 540
 Gly Ser Glu Thr Leu Lys Leu Gln Gln Leu Val Glu Asn Ile Asp Lys
 545 550 555 560
 Ala Thr Thr Asp Pro Asn Glu Cys Leu Ile Cys His Arg Val Leu Ser
 565 570 575
 Cys Gln Ser Ser Leu Lys Met His Tyr Arg Thr His Thr Gly Glu Arg
 580 585 590
 Pro Phe Gln Cys Lys Ile Cys Gly Arg Ala Phe Ser Thr Lys Gly Asn
 595 600 605
 Leu Lys Thr His Leu Gly Val His Arg Thr Asn Thr Ser Ile Lys Thr
 610 615 620
 Gln His Ser Cys Pro Ile Cys Gln Lys Lys Phe Thr Asn Ala Val Met
 625 630 635 640

[0031]

Leu Gln Gln His Ile Arg Met His Met Gly Gly Gln Ile Pro Asn Thr
 645 650 655
 Pro Leu Pro Glu Asn Pro Cys Asp Phe Thr Gly Ser Glu Pro Met Thr
 660 665 670
 Val Gly Glu Asn Gly Ser Thr Gly Ala Ile Cys His Asp Asp Val Ile
 675 680 685
 Glu Ser Ile Asp Val Glu Glu Val Ser Ser Gln Glu Ala Pro Ser Ser
 690 695 700
 Ser Ser Lys Val Pro Thr Pro Leu Pro Ser Ile His Ser Ala Ser Pro
 705 710 715 720
 Thr Leu Gly Phe Ala Met Met Ala Ser Leu Asp Ala Pro Gly Lys Val
 725 730 735
 Gly Pro Ala Pro Phe Asn Leu Gln Arg Gln Gly Ser Arg Glu Asn Gly
 740 745 750
 Ser Val Glu Ser Asp Gly Leu Thr Asn Asp Ser Ser Ser Leu Met Gly
 755 760 765
 Asp Gln Glu Tyr Gln Ser Arg Ser Pro Asp Ile Leu Glu Thr Thr Ser
 770 775 780
 Phe Gln Ala Leu Ser Pro Ala Asn Ser Gln Ala Glu Ser Ile Lys Ser
 785 790 795 800
 Lys Ser Pro Asp Ala Gly Ser Lys Ala Glu Ser Ser Glu Asn Ser Arg
 805 810 815
 Thr Glu Met Glu Gly Arg Ser Ser Leu Pro Ser Thr Phe Ile Arg Ala
 820 825 830
 Pro Pro Thr Tyr Val Lys Val Glu Val Pro Gly Thr Phe Val Gly Pro
 835 840 845
 Ser Thr Leu Ser Pro Gly Met Thr Pro Leu Leu Ala Ala Gln Pro Arg
 850 855 860
 Arg Gln Ala Lys Gln His Gly Cys Thr Arg Cys Gly Lys Asn Phe Ser
 865 870 875 880
 Ser Ala Ser Ala Leu Gln Ile His Glu Arg Thr His Thr Gly Glu Lys
 885 890 895
 Pro Phe Val Cys Asn Ile Cys Gly Arg Ala Phe Thr Thr Lys Gly Asn
 900 905 910
 Leu Lys Val His Tyr Met Thr His Gly Ala Asn Asn Asn Ser Ala Arg
 915 920 925
 Arg Gly Arg Lys Leu Ala Ile Glu Asn Thr Met Ala Leu Leu Gly Thr
 930 935 940
 Asp Gly Lys Arg Val Ser Glu Ile Phe Pro Lys Glu Ile Leu Ala Pro

[0032]

```

945          950          955          960
Ser Val Asn Val Asp Pro Val Val Trp Asn Gln Tyr Thr Ser Met Leu
          965          970          975

Asn Gly Gly Leu Ala Val Lys Thr Asn Glu Ile Ser Val Ile Gln Ser
          980          985          990

Gly Gly Val Pro Thr Leu Pro Val Ser Leu Gly Ala Thr Ser Val Val
          995          1000          1005

Asn Asn Ala Thr Val Ser Lys Met Asp Gly Ser Gln Ser Gly Ile
          1010          1015          1020

Ser Ala Asp Val Glu Lys Pro Ser Ala Thr Asp Gly Val Pro Lys
          1025          1030          1035

His Gln Phe Pro His Phe Leu Glu Glu Asn Lys Ile Ala Val Ser
          1040          1045          1050

<210> 12
<211> 616
<212> PRT
<213> 智人(Homo sapiens)

<220>
<221> 新的或少见的特征
<222> (113)..(113)
<223> Xaa可以是任何天然存在的氨基酸

<220>
<221> 新的或少见的特征
<222> (141)..(141)
<223> Xaa可以是任何天然存在的氨基酸

<220>
<221> 新的或少见的特征
<222> (442)..(442)
<223> Xaa可以是任何天然存在的氨基酸

<220>
<221> 新的或少见的特征
<222> (444)..(445)
<223> Xaa可以是任何天然存在的氨基酸

<220>
<221> 新的或少见的特征
<222> (456)..(456)
<223> Xaa可以是任何天然存在的氨基酸

<220>
<221> 新的或少见的特征
<222> (602)..(602)
<223> Xaa可以是任何天然存在的氨基酸

<220>
<221> 新的或少见的特征
<222> (613)..(613)
<223> Xaa可以是任何天然存在的氨基酸
<400> 12

Met Ser Arg Arg Lys Gln Ala Lys Pro Gln His Ile Asn Ser Glu Glu
1          5          10          15

Asp Gln Gly Glu Gln Gln Pro Gln Gln Thr Pro Glu Phe Ala Asp
20          25          30

Ala Ala Pro Ala Ala Pro Ala Ala Gly Glu Leu Gly Ala Pro Val Asn
35          40          45

```

[0033]

His Pro Gly Asn Asp Glu Val Ala Ser Glu Asp Glu Ala Thr Val Lys
 50 55 60
 Arg Leu Arg Arg Glu Glu Thr His Val Cys Glu Lys Cys Cys Ala Glu
 65 70 75 80
 Phe Phe Ser Ile Ser Glu Phe Leu Glu His Lys Lys Asn Cys Thr Lys
 85 90 95
 Asn Pro Pro Val Leu Ile Met Asn Asp Ser Glu Gly Pro Val Pro Ser
 100 105 110
 Xaa Asp Phe Ser Gly Ala Val Leu Ser His Gln Pro Thr Ser Pro Gly
 115 120 125
 Ser Glu Asp Cys His Arg Glu Asn Gly Gly Ser Ser Xaa Asp Ile Lys
 130 135 140
 Glu Lys Pro Asp Ala Glu Ser Val Val Tyr Leu Lys Thr Glu Thr Ala
 145 150 155 160
 Leu Pro Pro Thr Pro Gln Asp Ile Ser Tyr Leu Ala Lys Gly Lys Val
 165 170 175
 Ala Asn Thr Asn Val Thr Leu Gln Ala Leu Arg Gly Thr Lys Val Ala
 180 185 190
 Val Asn Gln Arg Ser Ala Asp Ala Leu Pro Ala Pro Val Pro Gly Ala
 195 200 205
 Asn Ser Ile Pro Trp Val Leu Glu Gln Ile Leu Cys Leu Gln Gln Gln
 210 215 220
 Gln Leu Gln Gln Ile Gln Leu Thr Glu Gln Ile Arg Ile Gln Val Asn
 225 230 235 240
 Met Trp Ala Ser His Ala Leu His Ser Ser Gly Ala Gly Ala Asp Thr
 245 250 255
 Leu Lys Thr Leu Gly Ser His Met Ser Gln Gln Val Ser Ala Ala Val
 260 265 270
 Ala Leu Leu Ser Gln Lys Ala Gly Ser Gln Gly Leu Ser Leu Asp Ala
 275 280 285
 Leu Lys Gln Ala Lys Leu Pro His Ala Asn Ile Pro Ser Ala Thr Ser
 290 295 300
 Ser Leu Ser Pro Gly Leu Ala Pro Phe Thr Leu Lys Pro Asp Gly Thr
 305 310 315 320
 Arg Val Leu Pro Asn Val Met Ser Arg Leu Pro Ser Ala Leu Leu Pro
 325 330 335
 Gln Ala Pro Gly Ser Val Leu Phe Gln Ser Pro Phe Ser Thr Val Ala
 340 345 350
 Leu Asp Thr Ser Lys Lys Gly Lys Gly Lys Pro Pro Asn Ile Ser Ala

[0034]

355	360	365
Val Asp Val Lys Pro Lys Asp Glu Ala Ala Leu Tyr Lys His Lys Cys 370 375 380		
Arg Ser Ser Leu Pro Ser Thr Phe Ile Arg Ala Pro Pro Thr Tyr Val 385 390 395 400		
Lys Val Glu Val Pro Gly Thr Phe Val Gly Pro Ser Thr Leu Ser Pro 405 410 415		
Gly Met Thr Pro Leu Leu Ala Ala Gln Pro Arg Gly Gln Ala Lys Gln 420 425 430		
His Gly Cys Thr Arg Cys Gly Lys Asn Xaa Ser Xaa Xaa Ser Ala Leu 435 440 445		
Gln Ile His Glu Arg Thr His Xaa Gly Glu Lys Pro Phe Val Cys Asn 450 455 460		
Ile Cys Gly Arg Ala Phe Thr Thr Lys Gly Asn Leu Lys Val His Tyr 465 470 475 480		
Met Thr His Gly Ala Asn Asn Asn Ser Ala Arg Arg Gly Arg Lys Leu 485 490 495		
Ala Ile Glu Asn Thr Met Ala Leu Leu Gly Thr Asp Gly Lys Arg Val 500 505 510		
Ser Glu Ile Phe Pro Lys Glu Ile Leu Ala Pro Ser Val Asn Val Asp 515 520 525		
Pro Val Val Trp Asn Gln Tyr Thr Ser Met Leu Asn Gly Gly Leu Ala 530 535 540		
Val Lys Thr Asn Glu Ile Ser Val Ile Gln Ser Gly Gly Val Pro Thr 545 550 555 560		
Leu Pro Val Ser Leu Gly Ala Thr Ser Val Val Asn Asn Ala Thr Val 565 570 575		
Ser Lys Met Asp Gly Ser Gln Ser Gly Ile Ser Ala Asp Val Glu Lys 580 585 590		
Pro Ser Ala Thr Asp Gly Val Pro Lys Xaa Gln Phe Pro His Phe Leu 595 600 605		
Glu Glu Asn Lys Xaa Ala Val Ser 610 615		

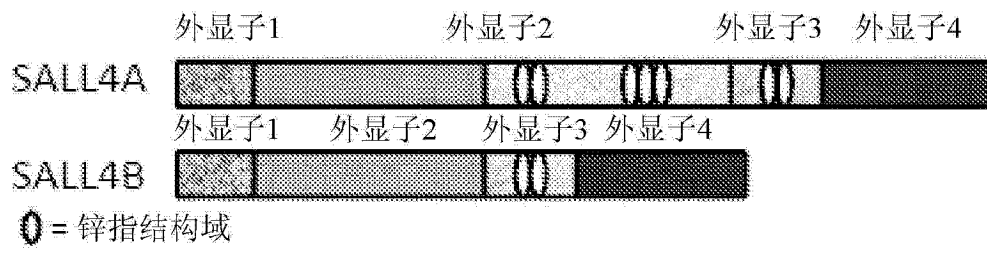


图 1

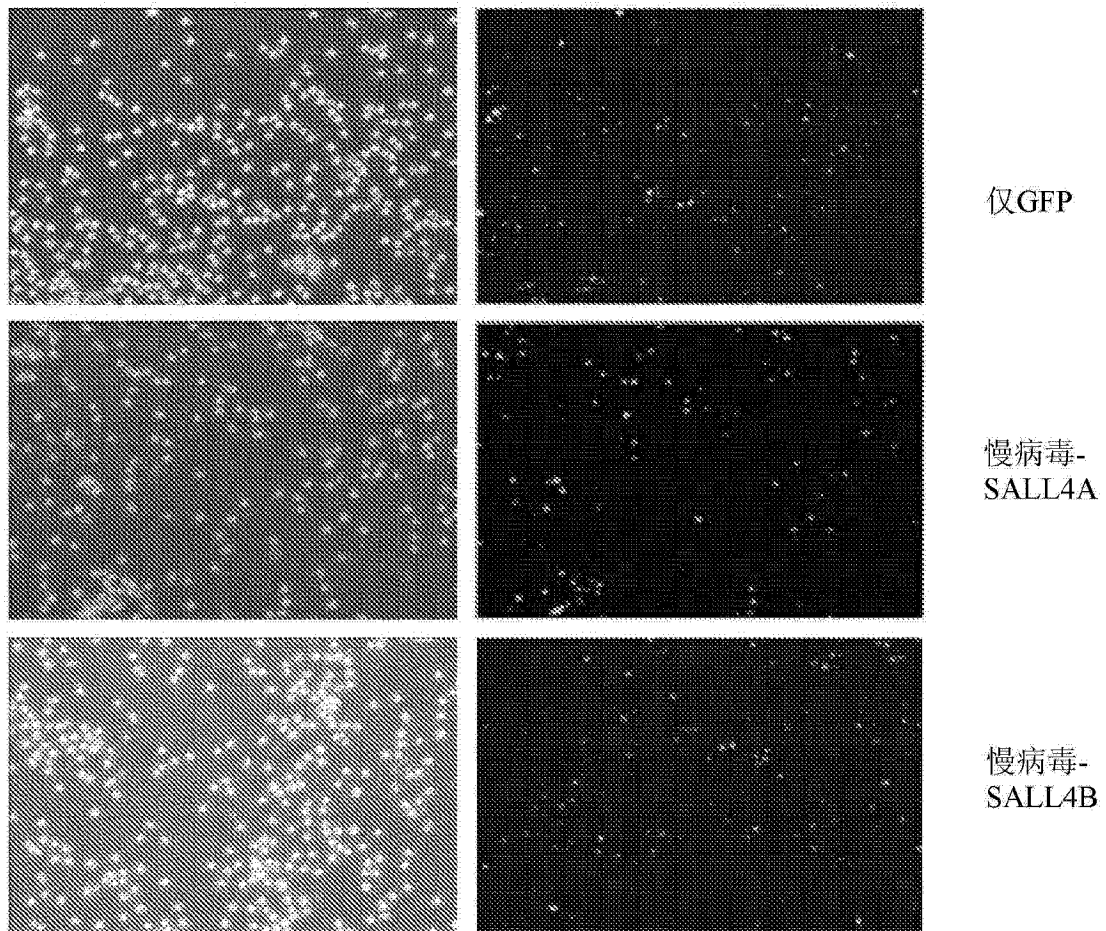


图 2

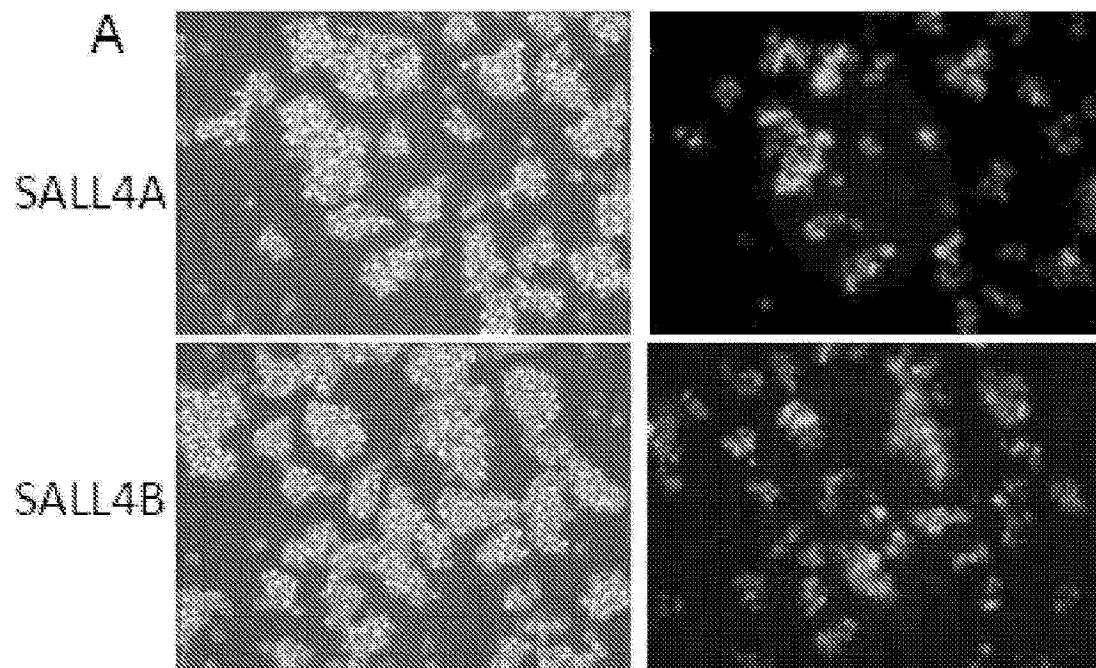
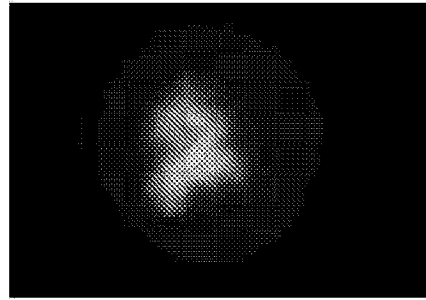
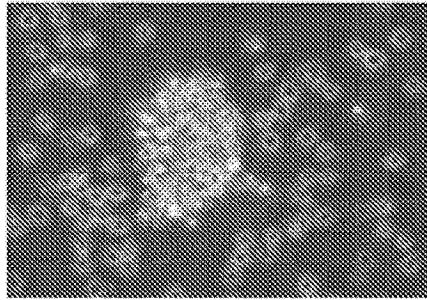


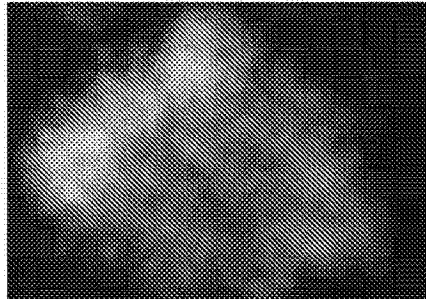
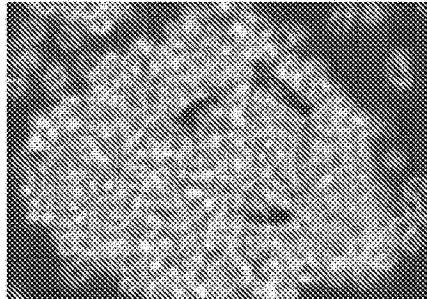
图 3

B

患者1



患者2



患者3

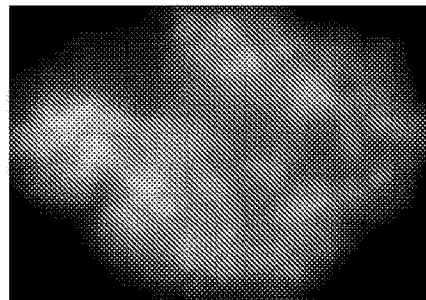
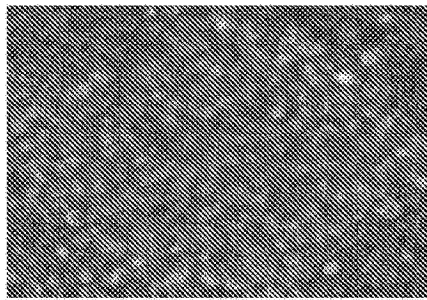


图 3(续)

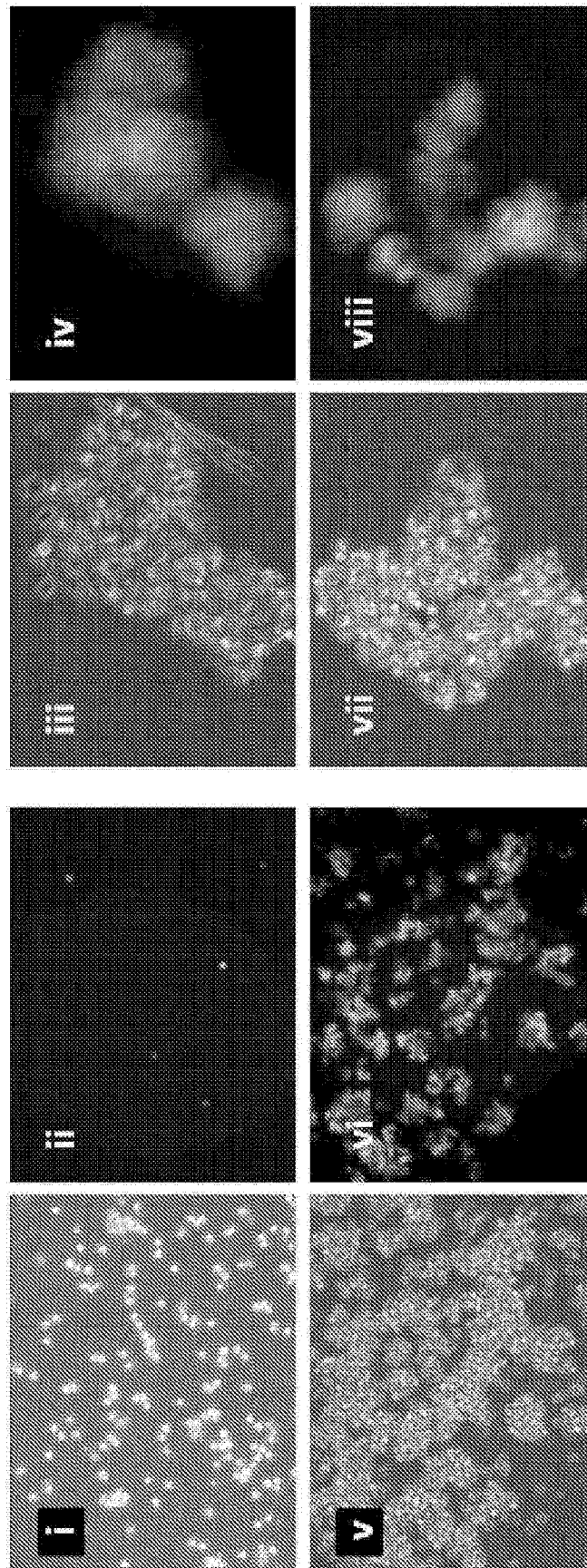


图 4

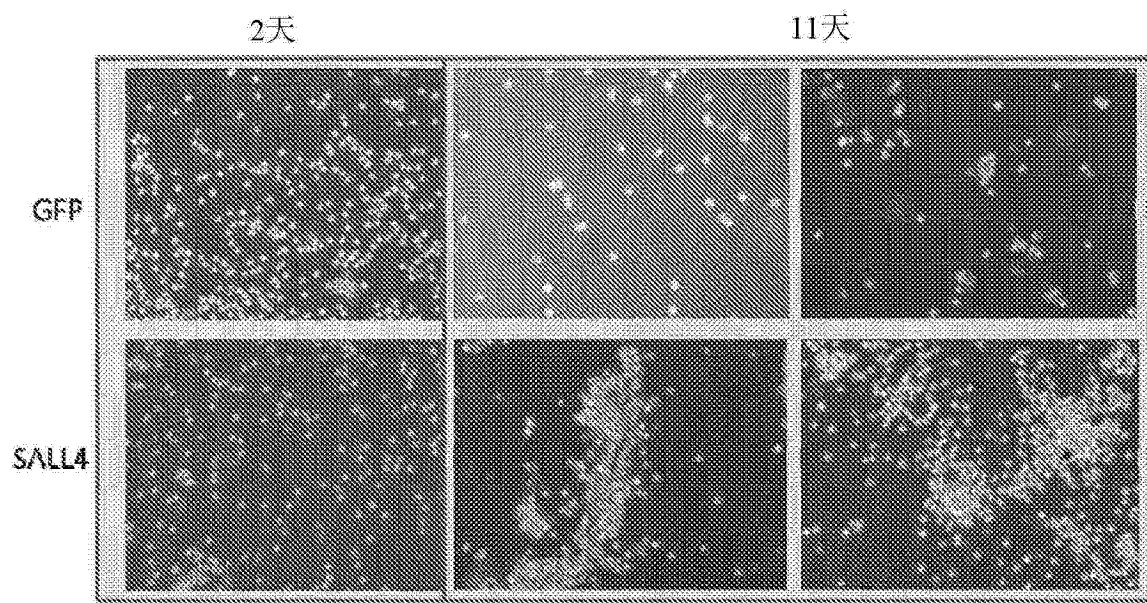


图 5

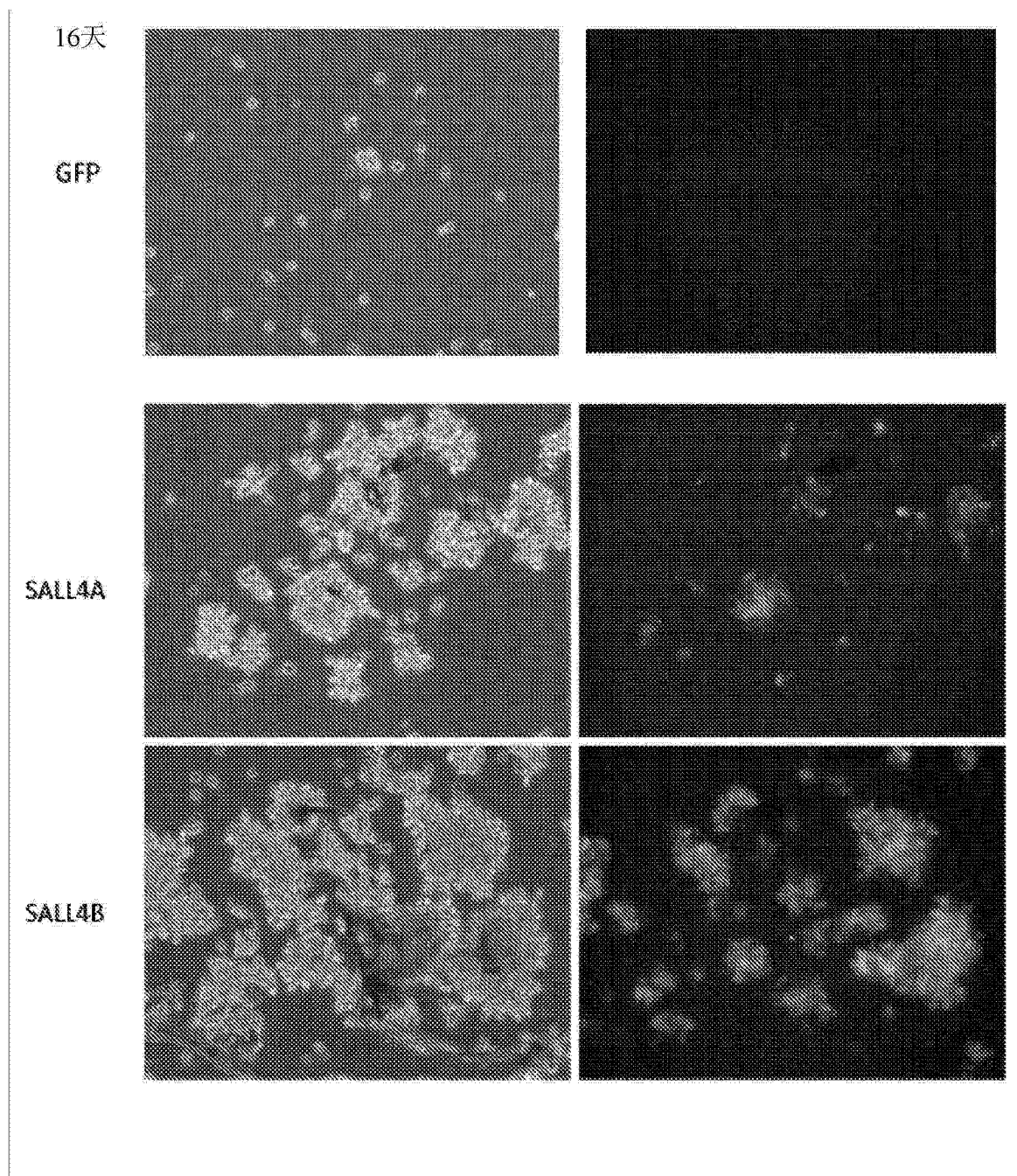


图 6

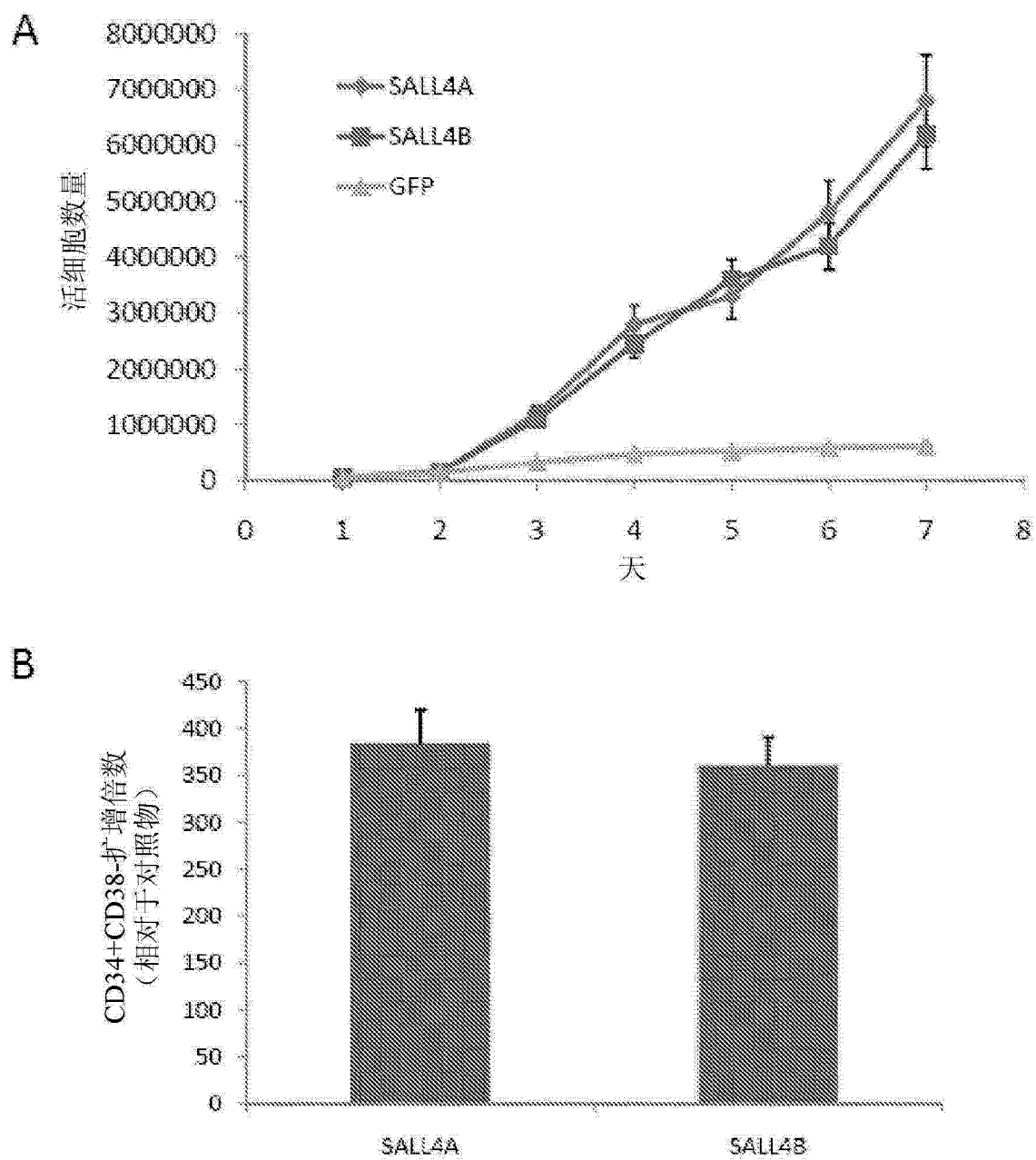


图 7

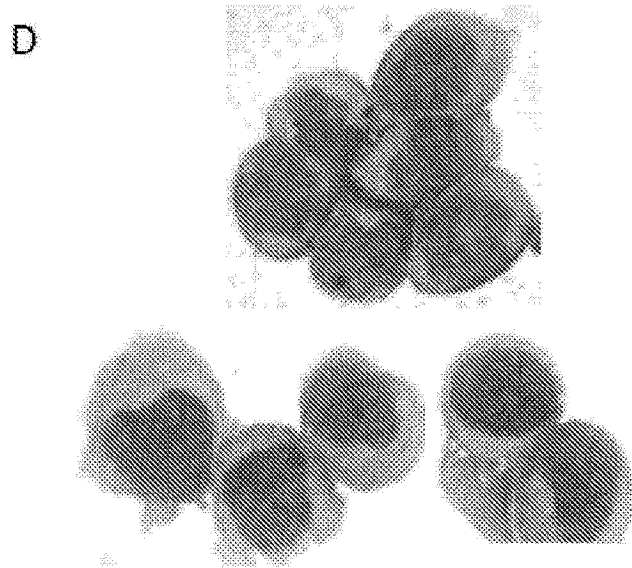
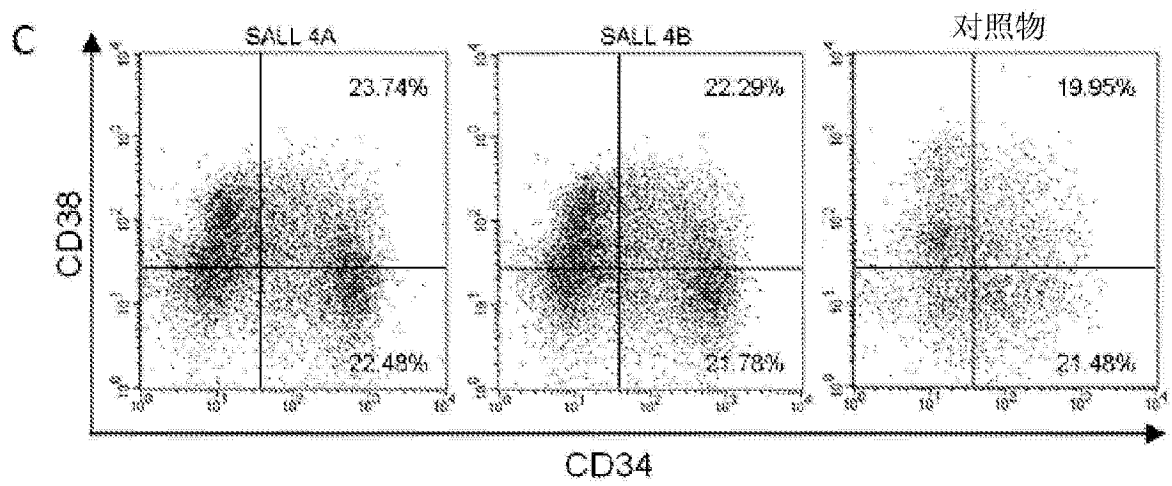


图7(续)

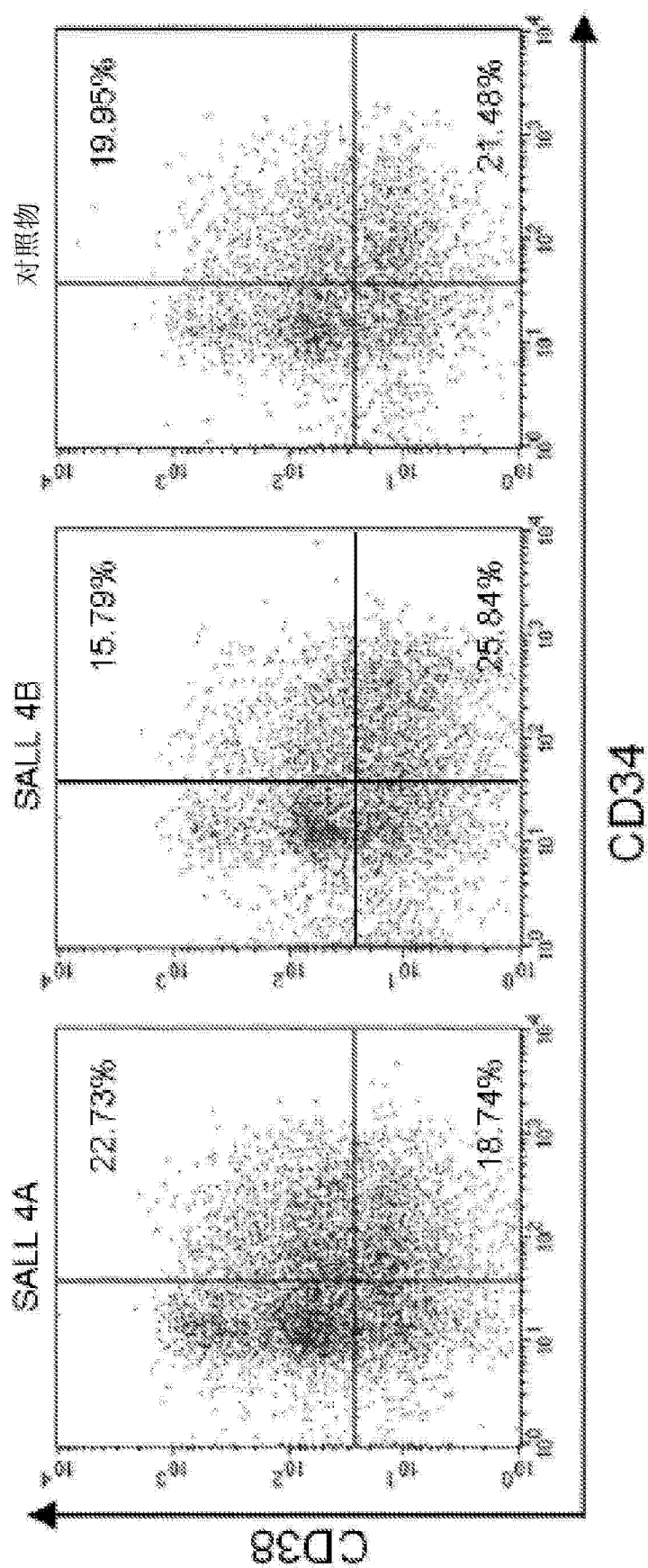


图 8

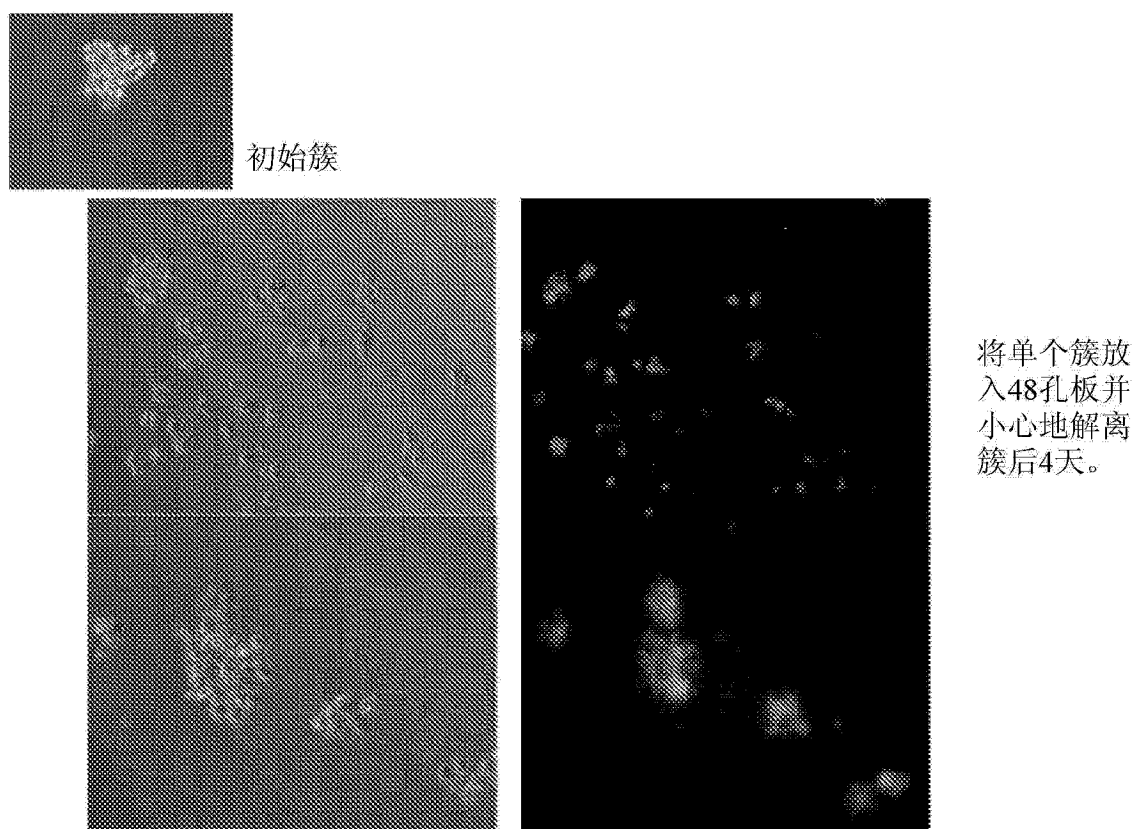


图 9



图 10

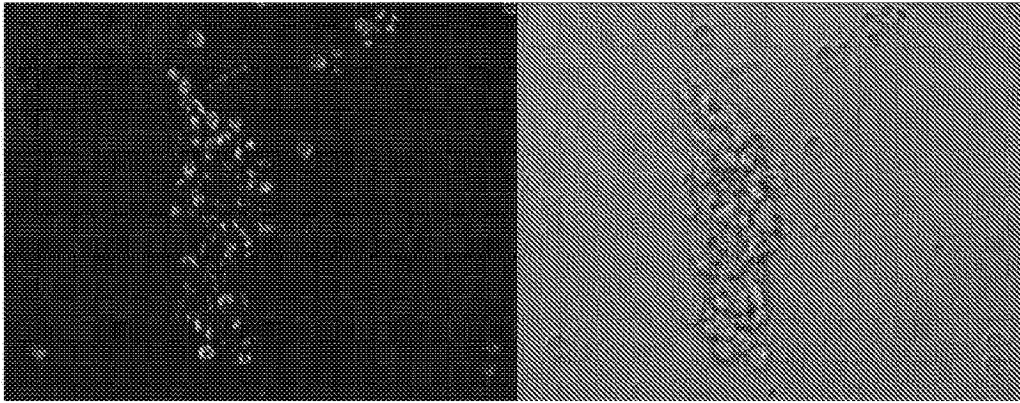


图 11

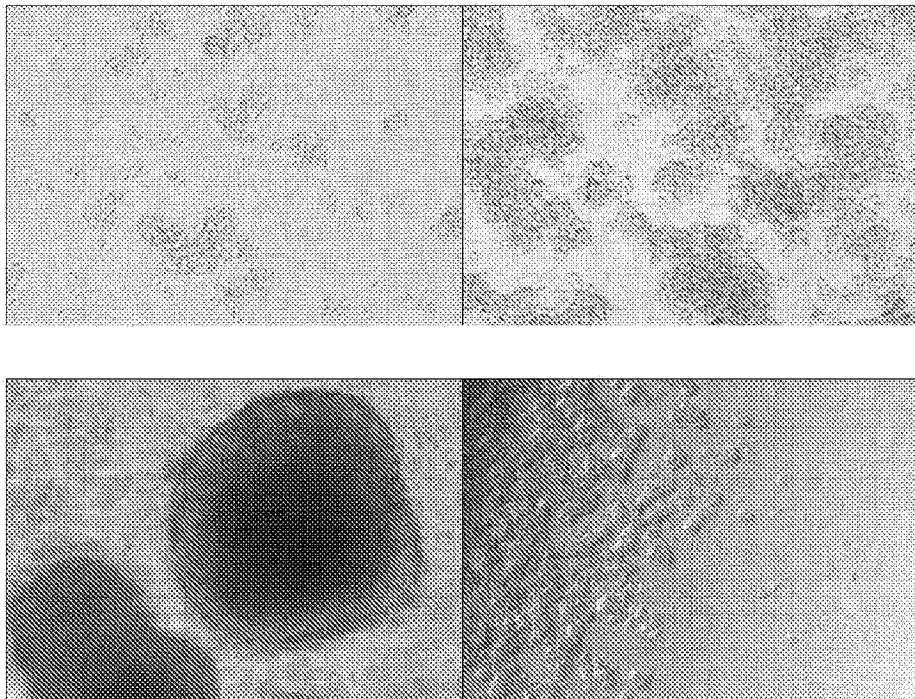


图 12

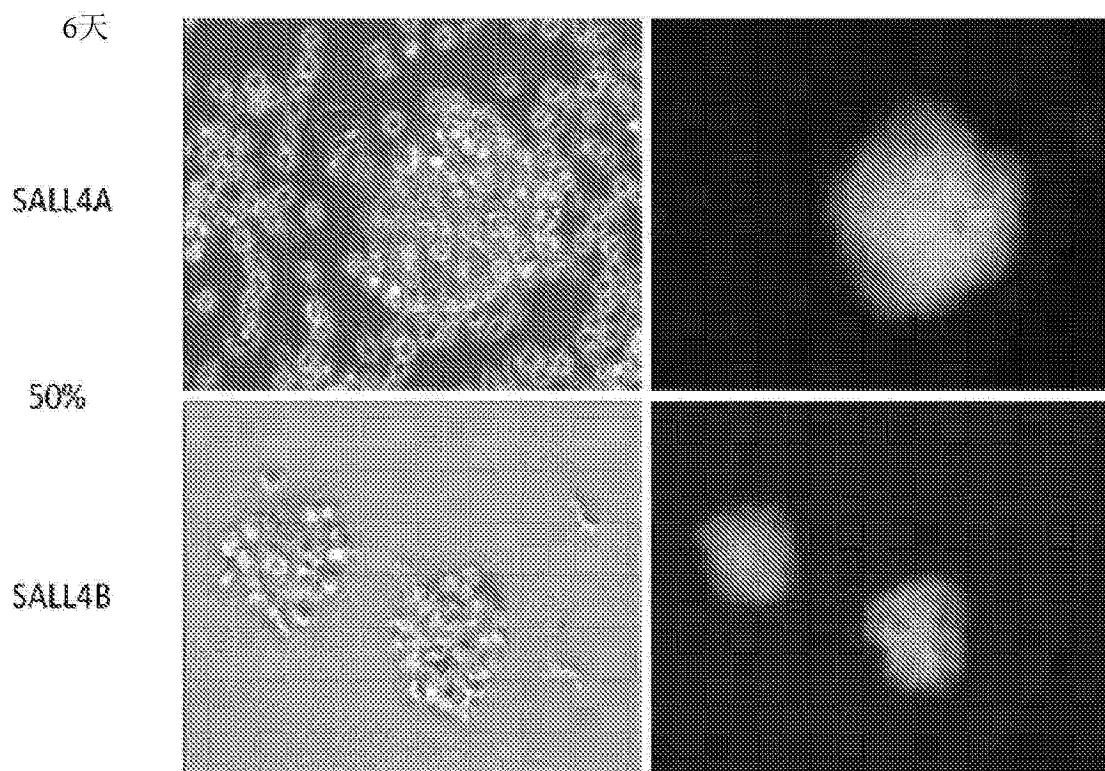


图 13

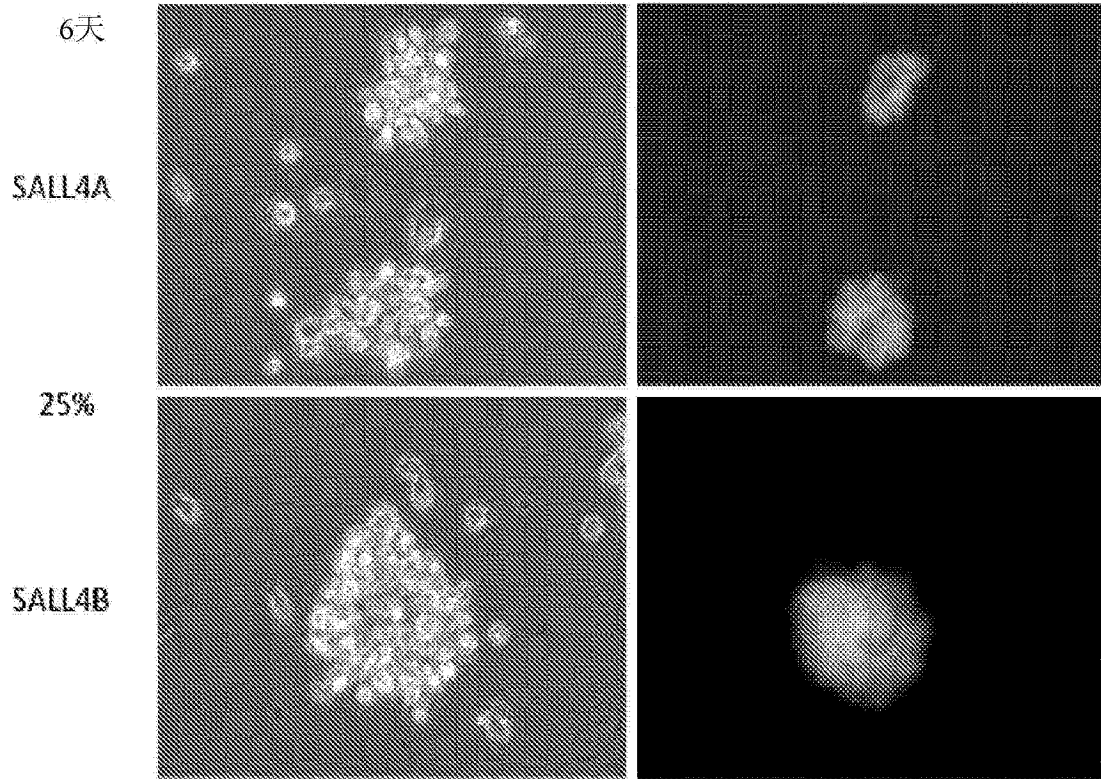


图 14

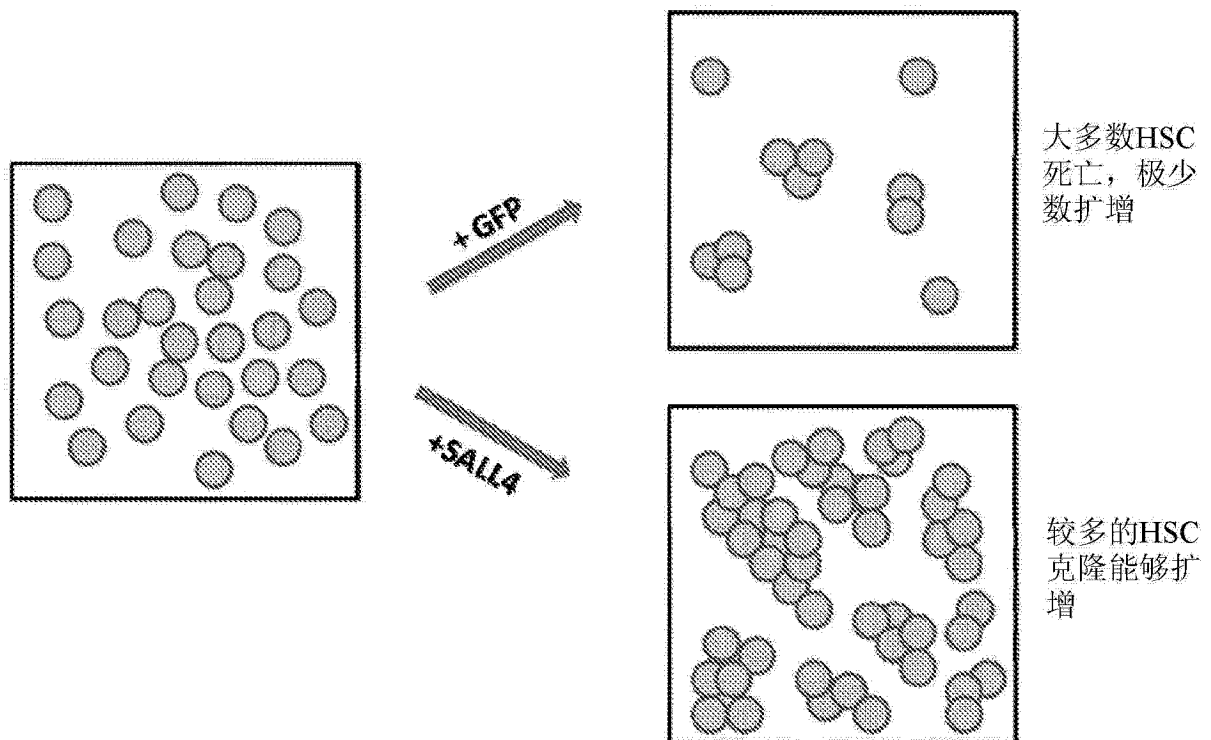


图 15

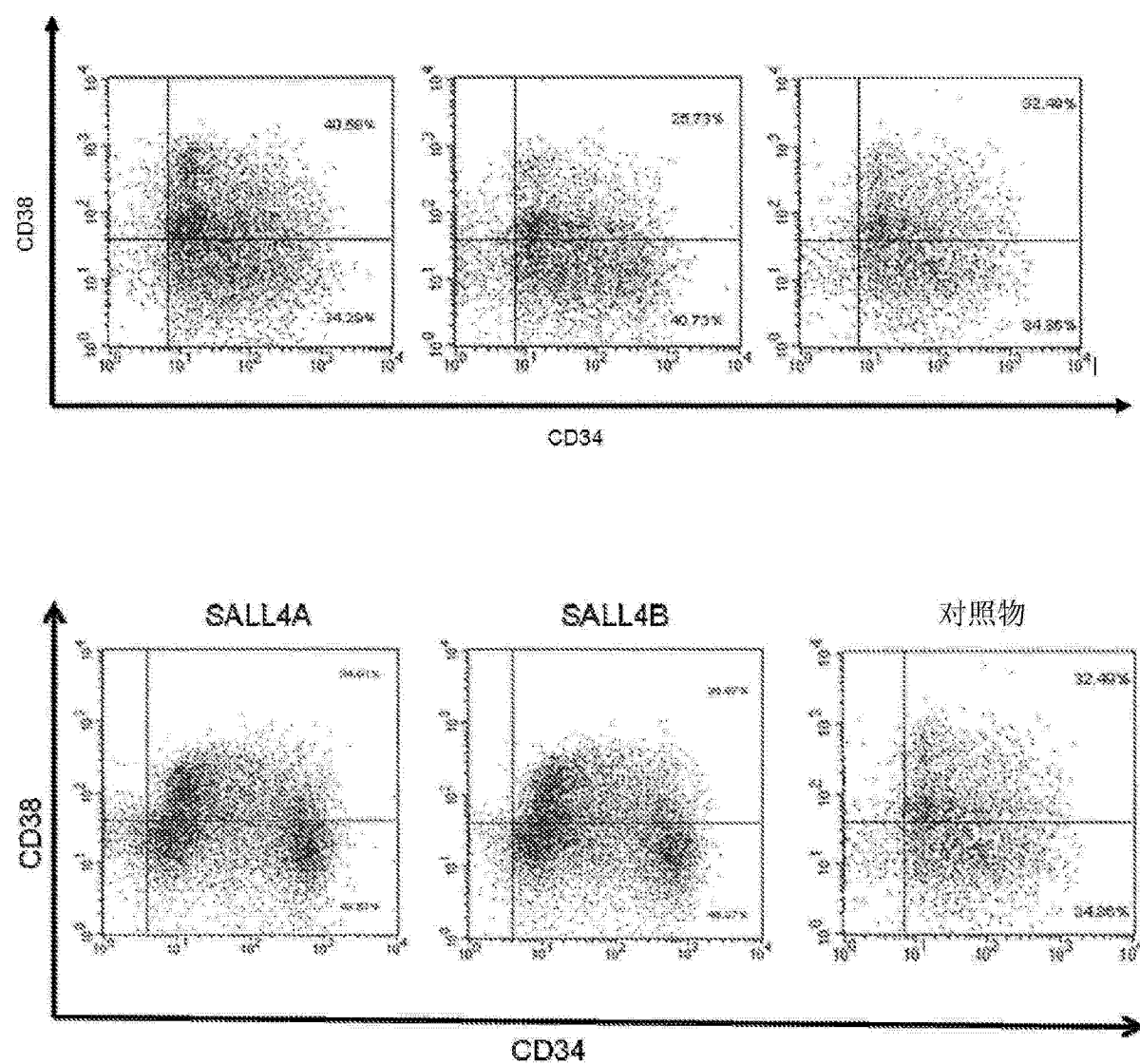


图 16

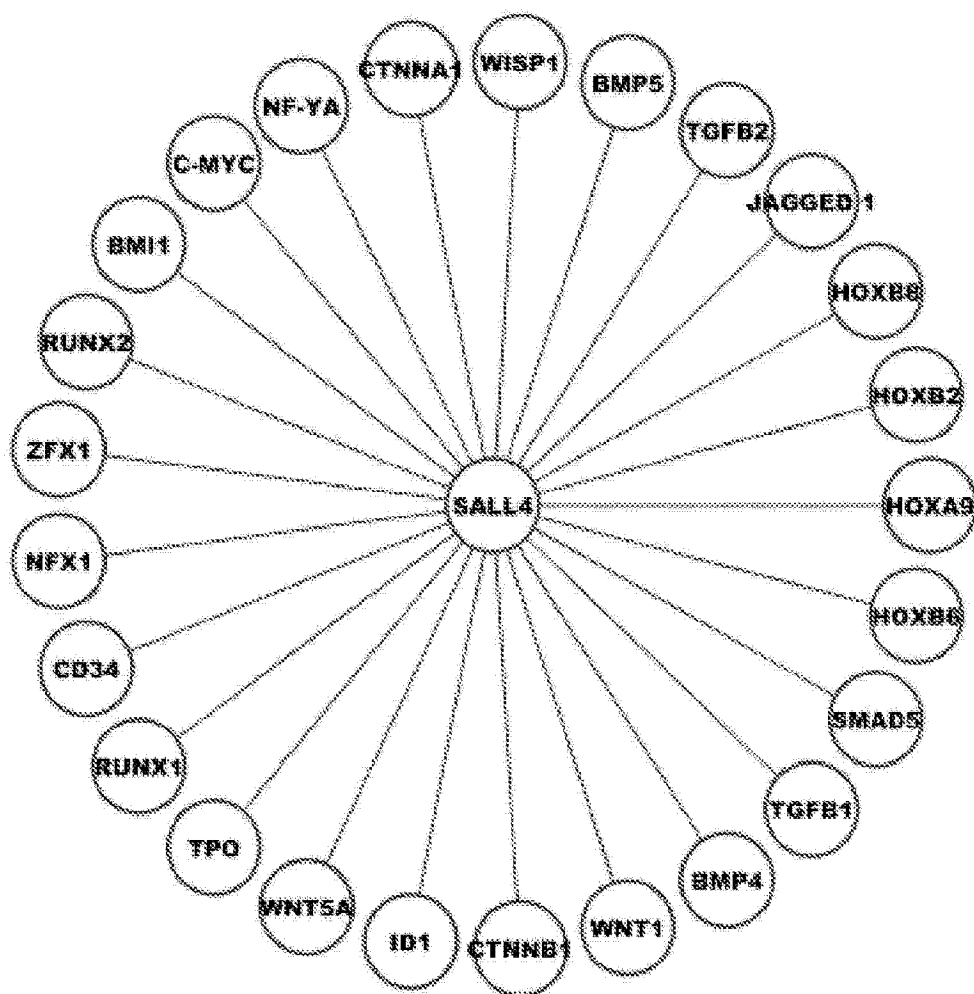


图 17

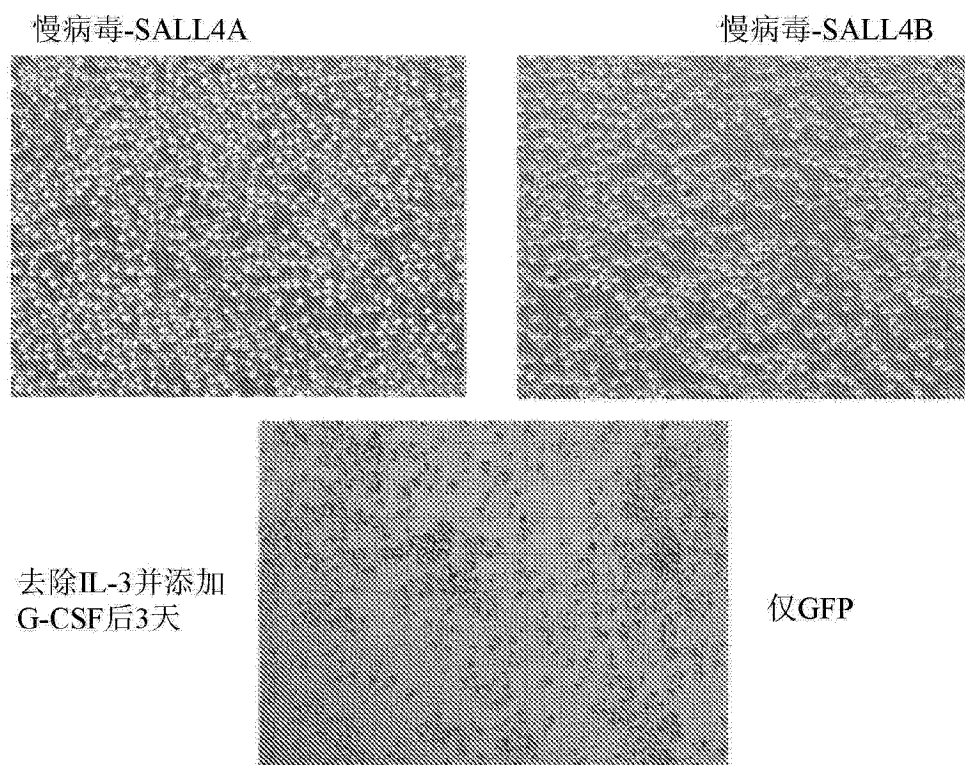


图 18A

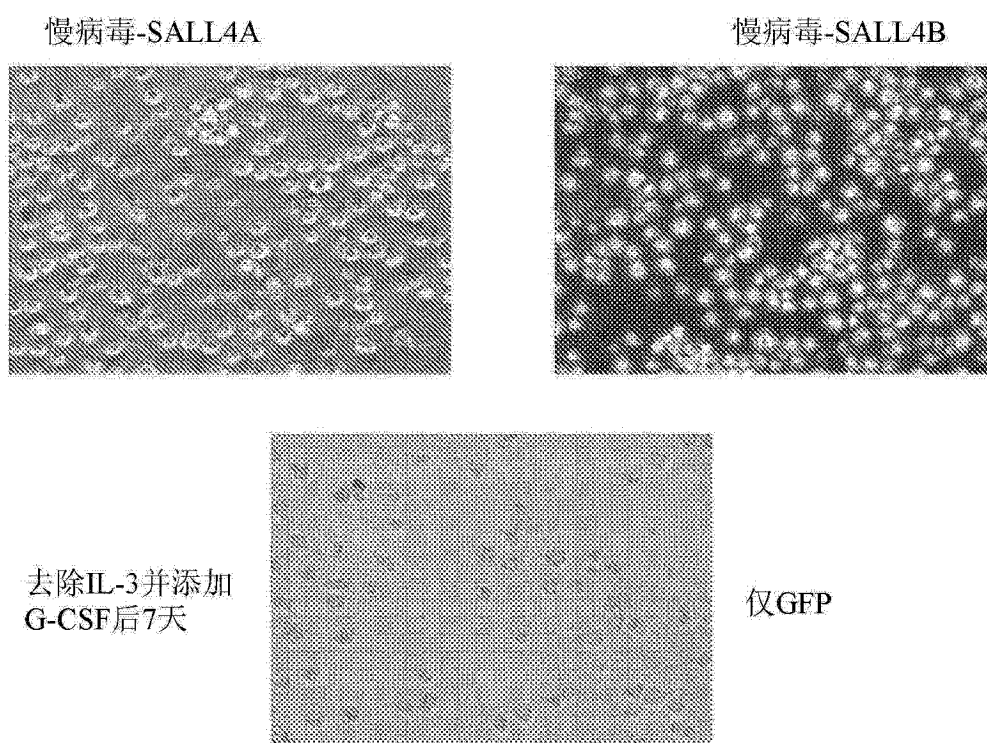


图 18B

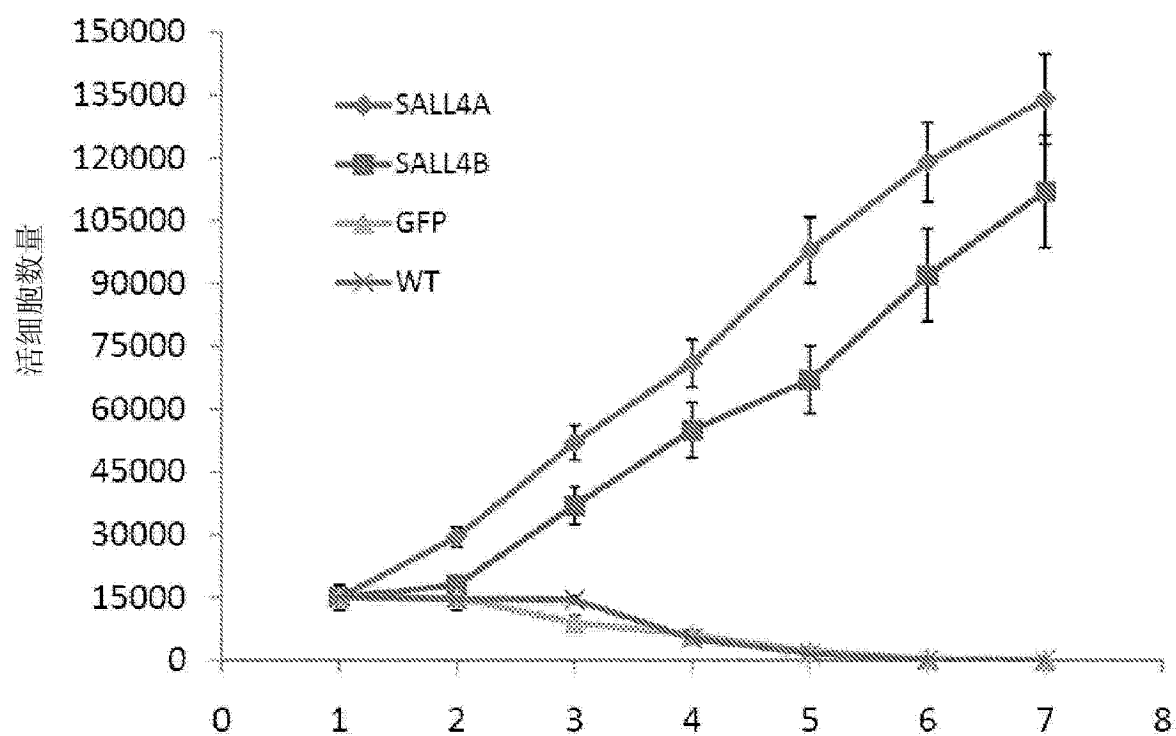


图 19

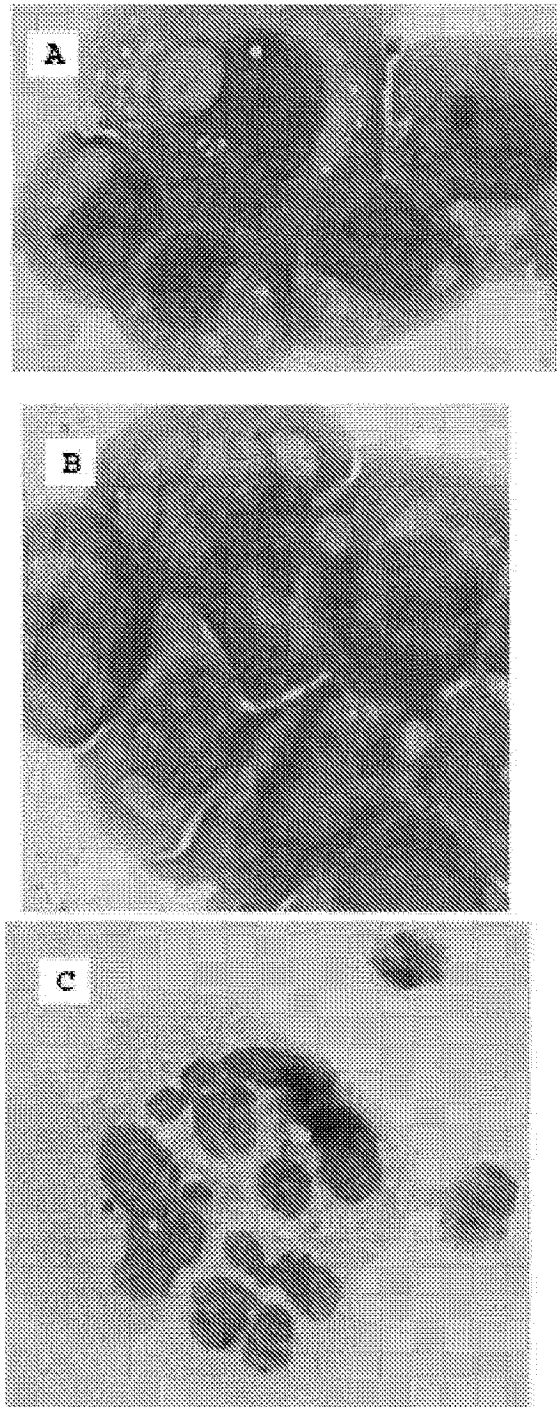


图 20

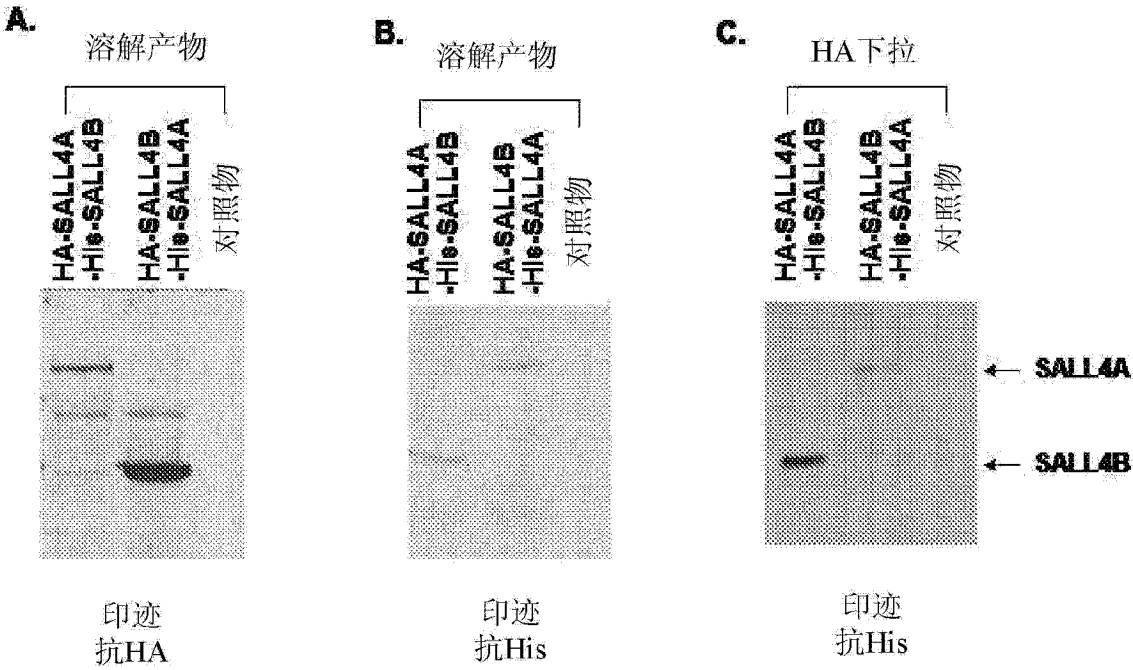
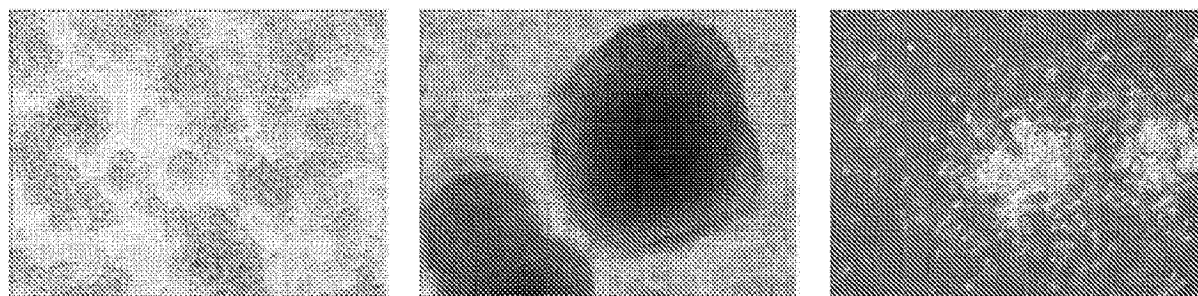
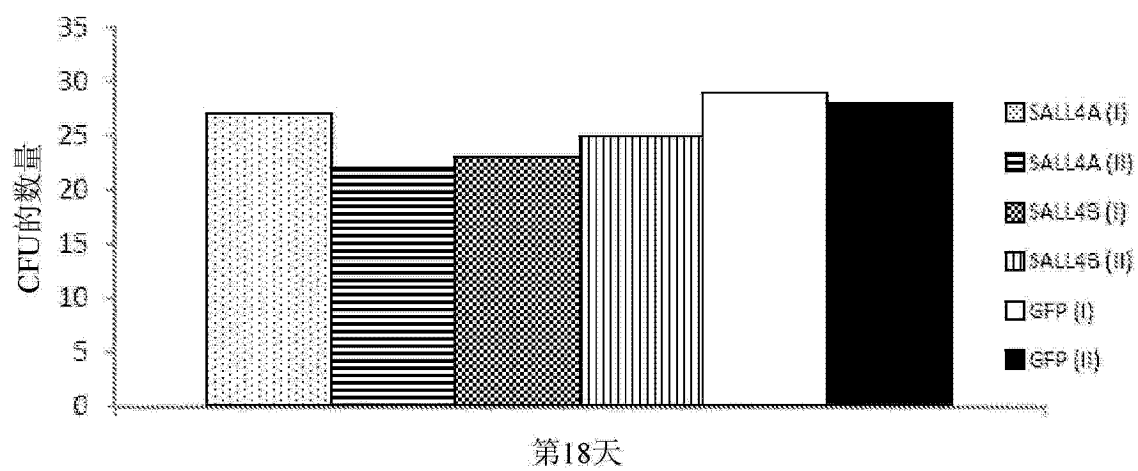


图 21

A



B



C

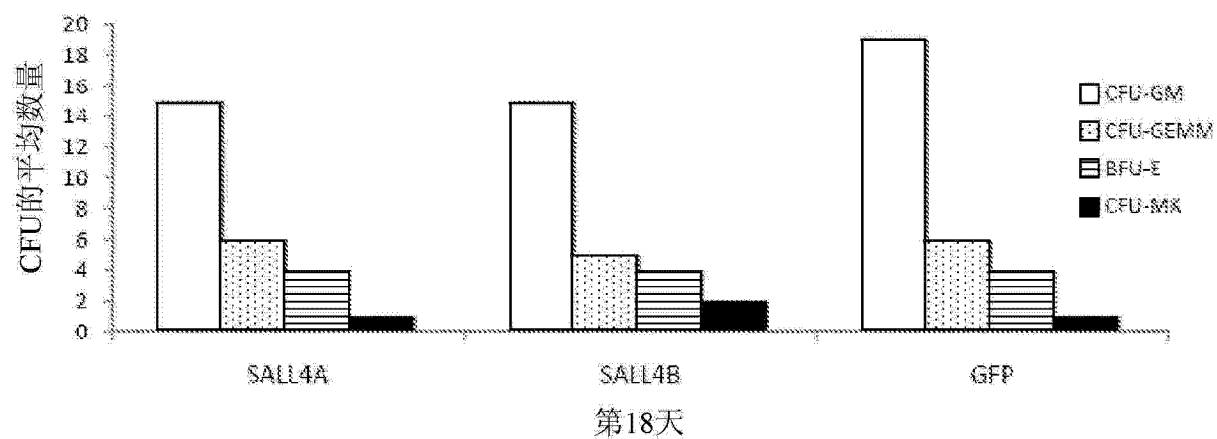


图 22

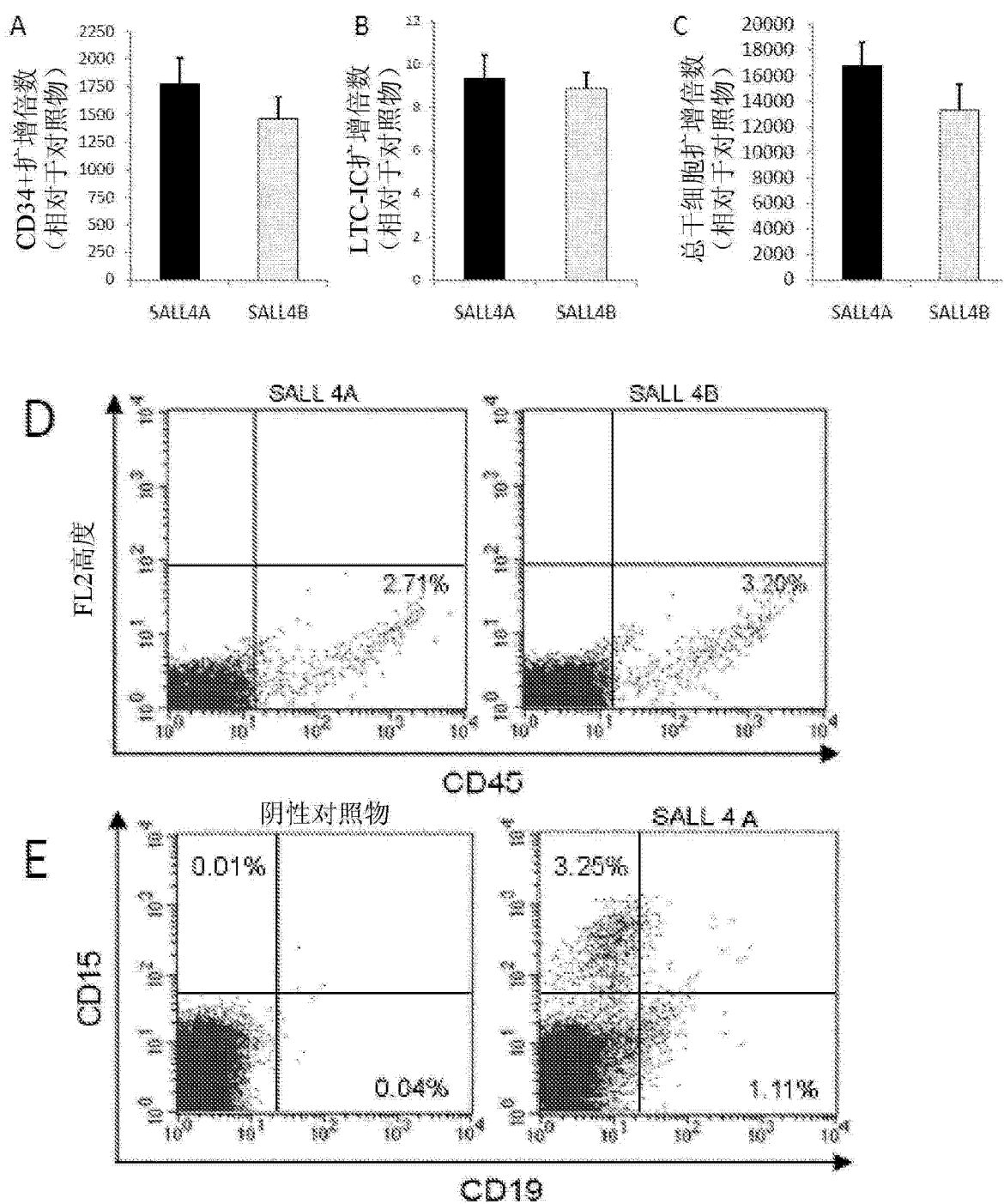


图 23

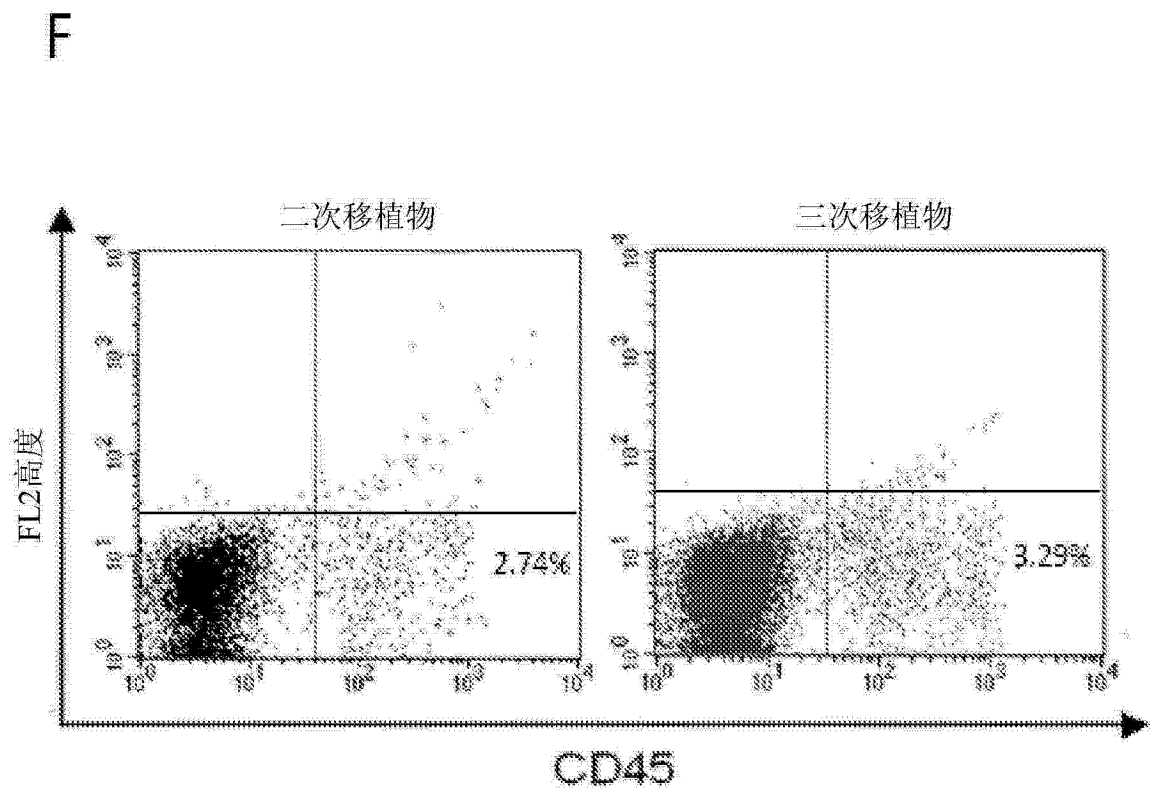


图 23(续)

G

三次移植谱系

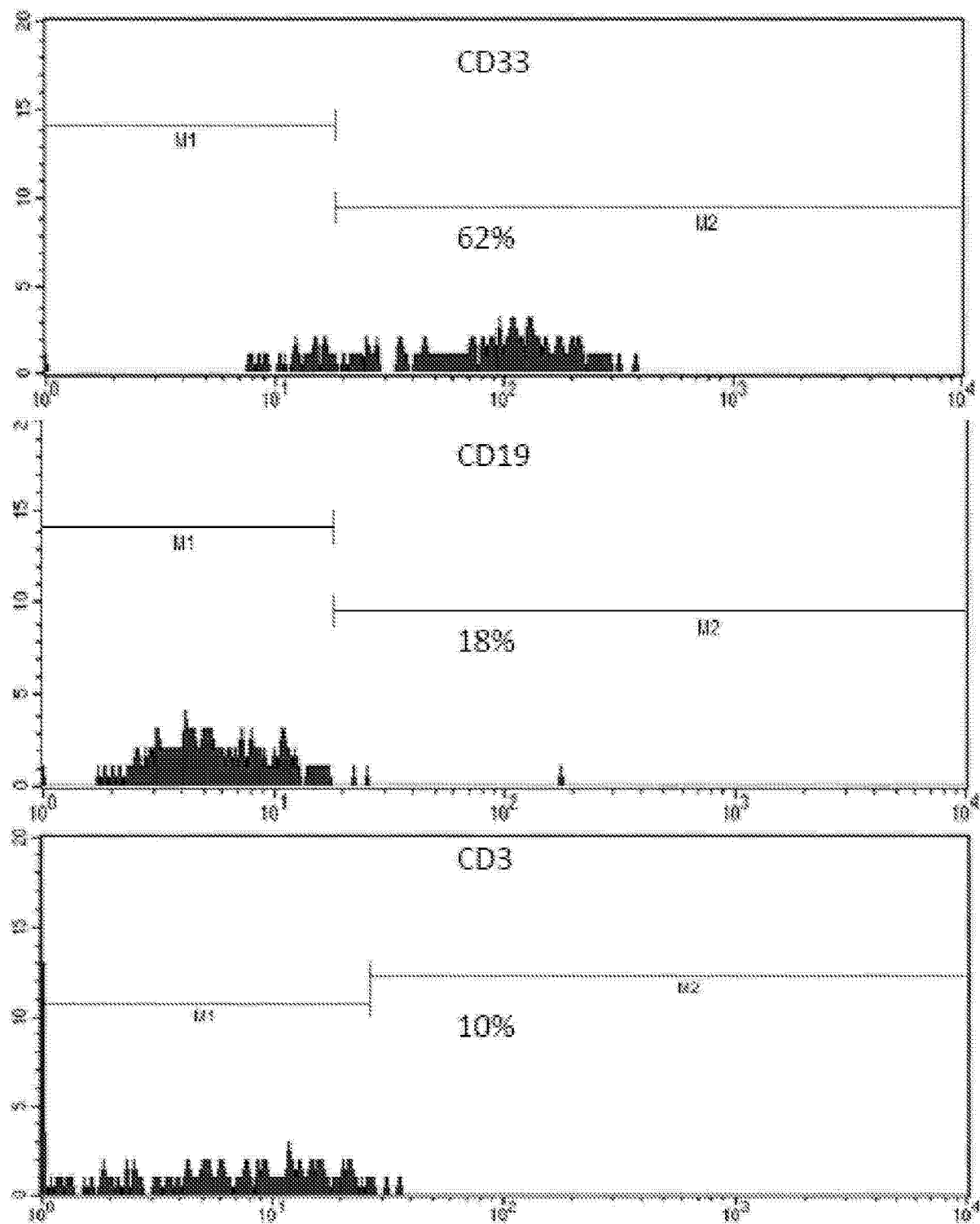


图 23(续)

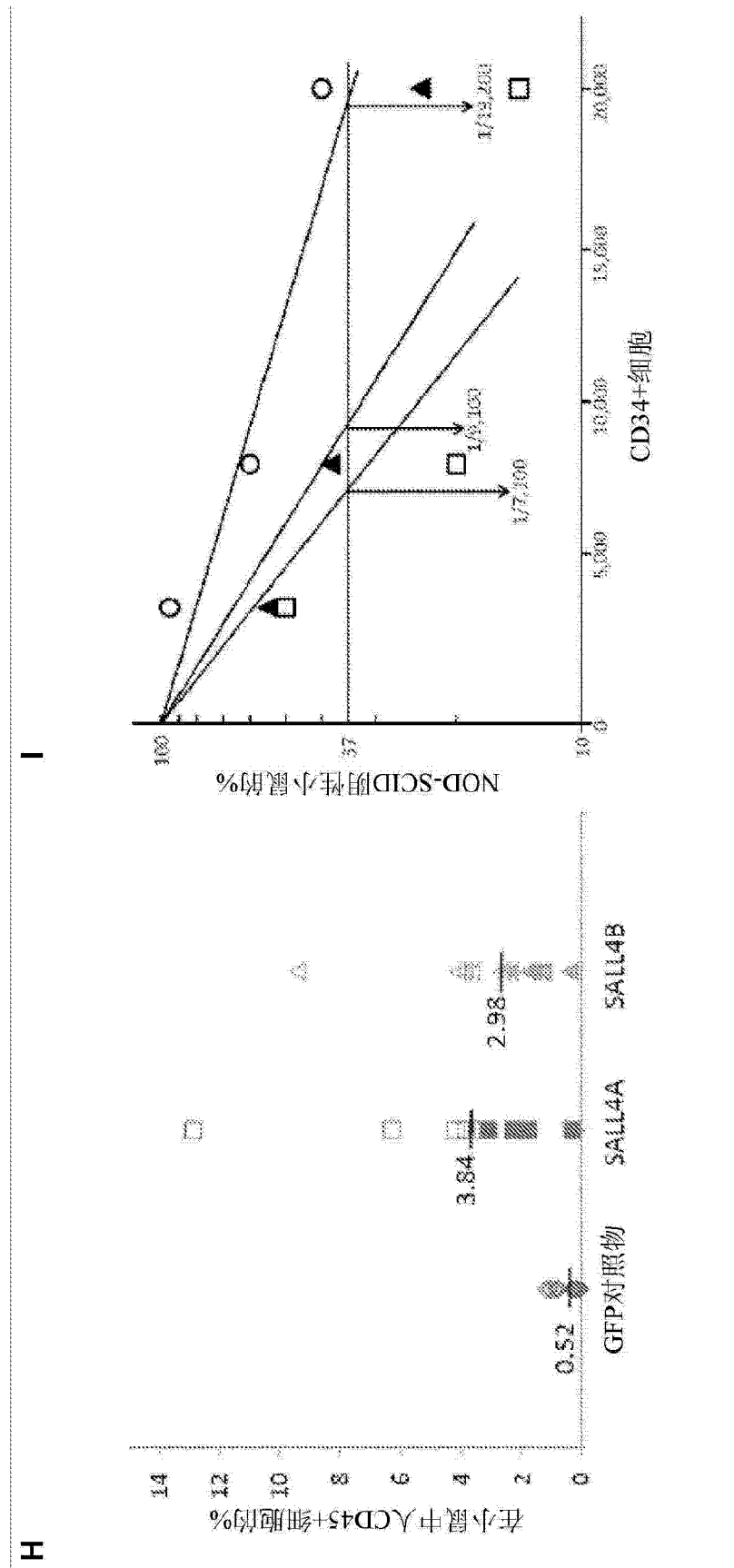


图 23(续)

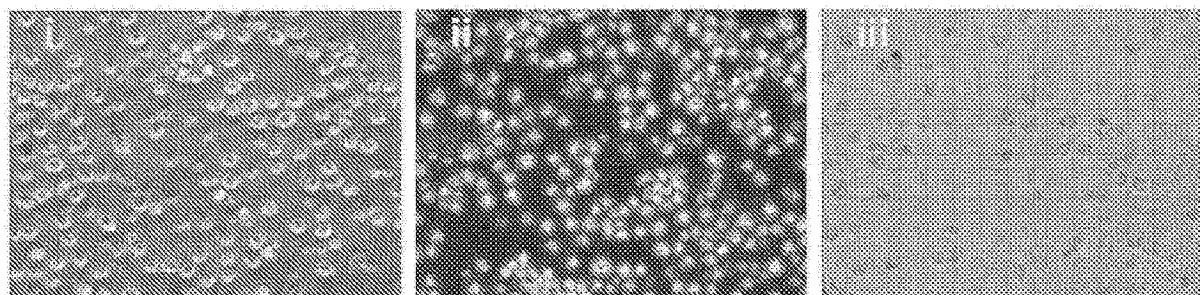


图 24

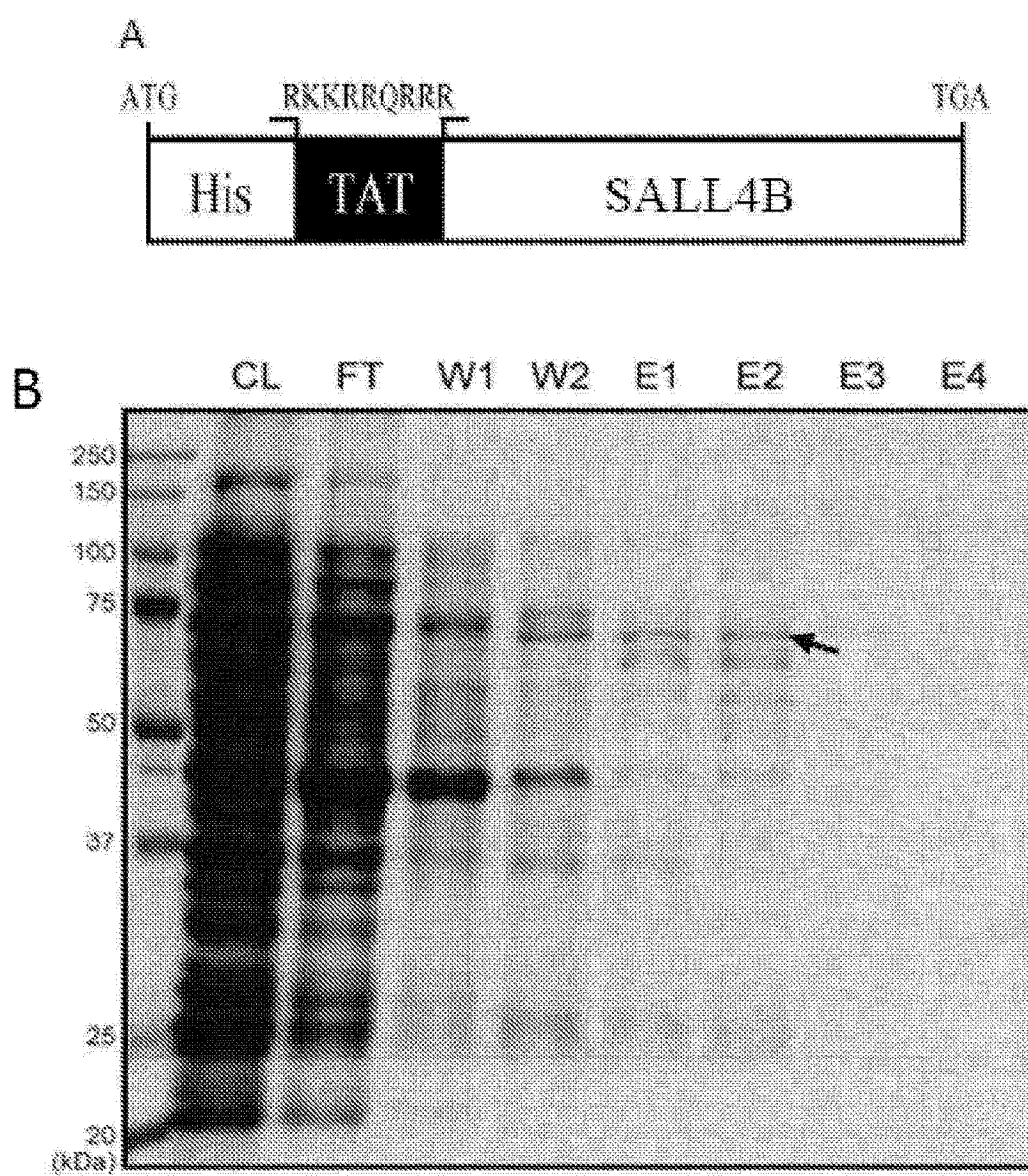


图 25

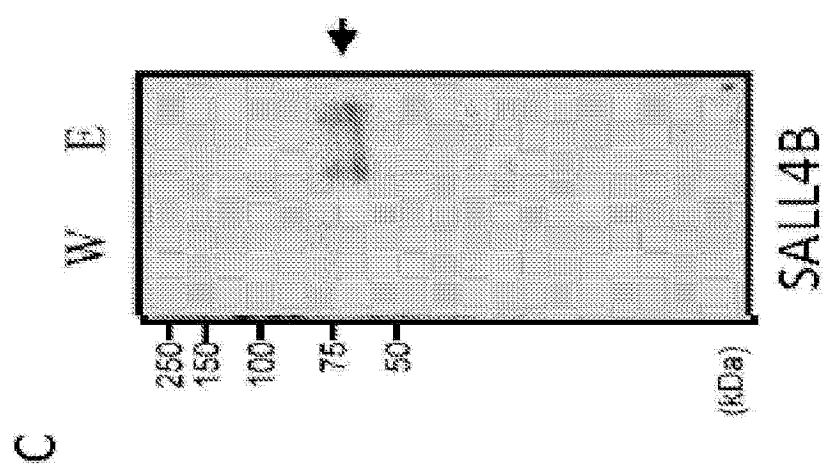


图 25(续)

D
 1 HSPRKEQAKPQ HINSEEDQSE QDPQQQTPEF ADAAPAAPAA GELGAPVNH
 51 GNDEVASEDE ATVKRLPREE THVCEKCCAE FFSISEFLEH KRNCTINPPV
 101 LIMNDEGCV FSEDFGAVL SHOPTSGSK DCHRENQSS EDNKEKPD
 151 SVVYLKETA LPPTPDISY LAKKVANTM VTLOALRGTK VAVHQSADA
 201 LPAPVPGAN IPVULEQILC LQQQLQCIQ LTEQIRIQVN NWASHALHSS
 251 GAGADTLKTL GSHMSQWSA AVALLSQKAG SQGLSLDALK QAKLPHANIP
 301 SATSSLSPGL APFTLKPDGT RVLPNVMSRL PSALLPQAPG SVLFQSPFST
 351 VALDTSKKCK GKPPNISAVD VPKDEAALY KKKCKKCSKV FQTPSSIQIH
 401 LRSHTGERPF VCSNCGHRT TKGNLKVHFN KHPQVKAAPQ LFAEFQDQVA
 451 AGNGIPYALS VPDPIDEPSL ELDSKPVLYT TSVGLPQNL SGTMPKDLTG
 501 GSLPGDLQFG PSPEZEGOPT LRGVGPNTMS PRAGGFQCSG TPEFGSETLK
 551 LQQLVENIDK ATTDPNELI CHRVLSQSS LKHNYKTHG KRPQCKICG
 601 RAFTTKGNLK THLGVHRTNT SHTQHCPI GQKFTNAVH LQQRIRNHNG
 651 GQTFNTPLPE NPCDFTCSGP MTVGNCSTG AICHDDVIES IDVKEVESQ
 701 APSSSSKVPF PLPSIHASAP TLOFANHSL DAPGVGPAP FNLONGGSR
 751 MCSVESDOLT NDSLSLGDQ EYQSRSPDL ETTFQALSP ANSQAESIES
 801 KSPDAGSKAE SSENSPTENE GSSLPSTFI RAPPYVUVE VPGTFVGPST
 851 LSPGMTPLLA AOPRRQAKQH CTRCGNFS SASALQIHER THTGKPFVC
 901 NICGRAFTIK GNKVVHYTH GANNSEKRC KKLAIENTHA LLGTDGKRVN
 951 EIFFKELILAP EVNVDFVVM OTTEHLNGL AVKINEISVI QSCGVPTLW
 1001 SLGATSVVNN ATVSKMDGSQ SGISADVEKP SATDGVPKHQ FPHFLEENK
 1051 KVS

由LC-MS/MS鉴定的具有His-TAT-SALL4B的肽呈黑色字体

图 25(续)

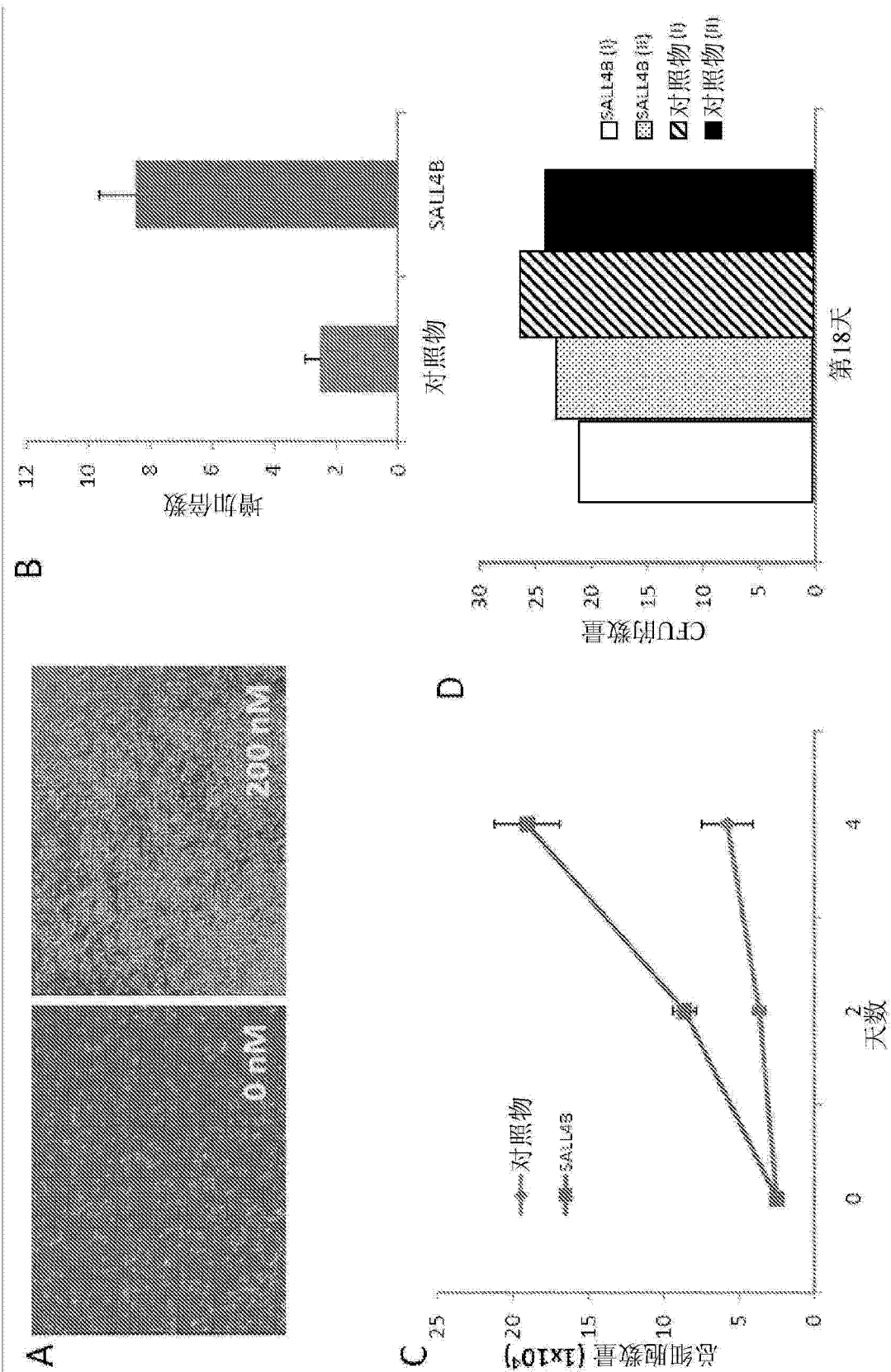


图 26

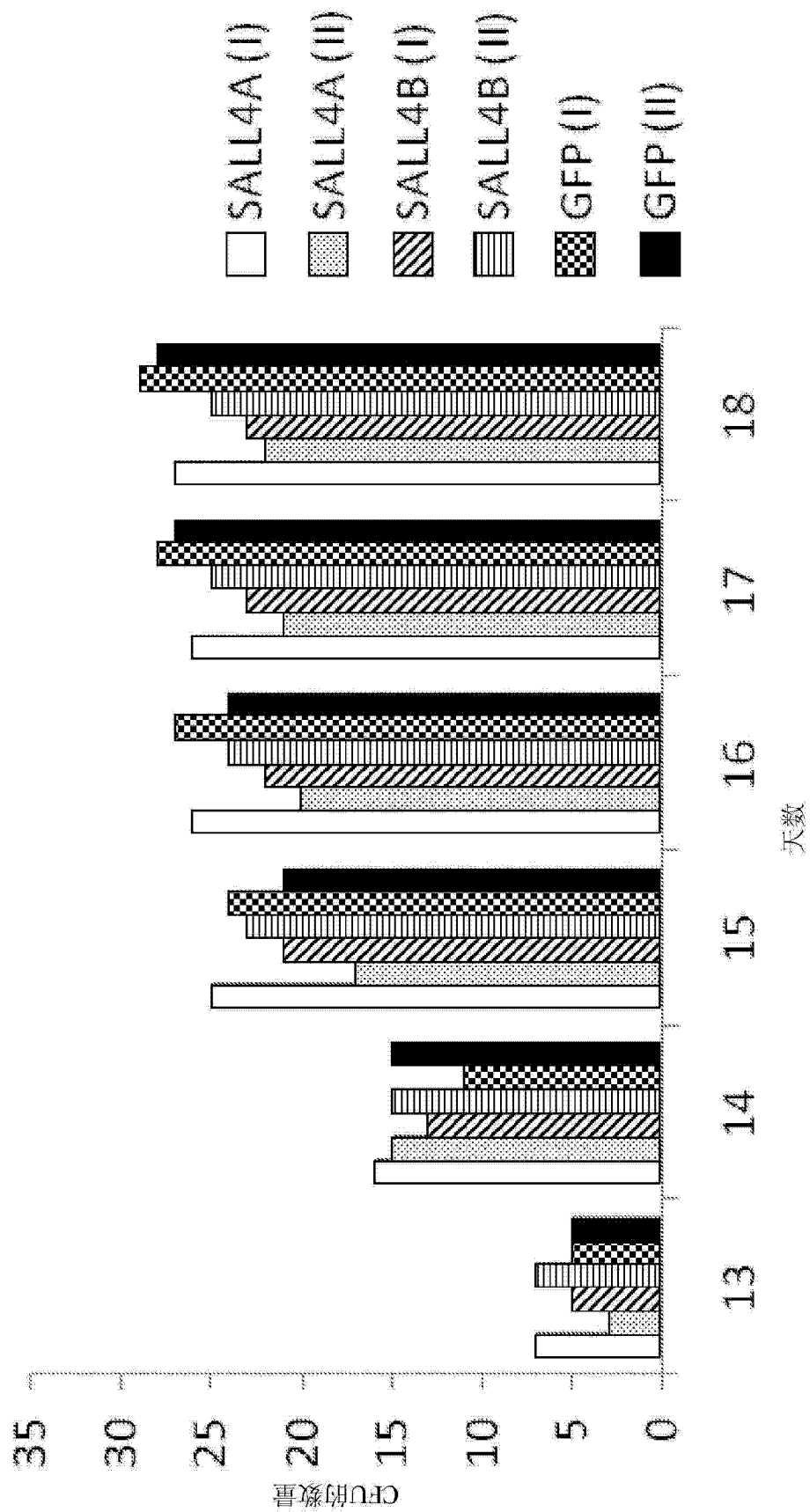


图 27A

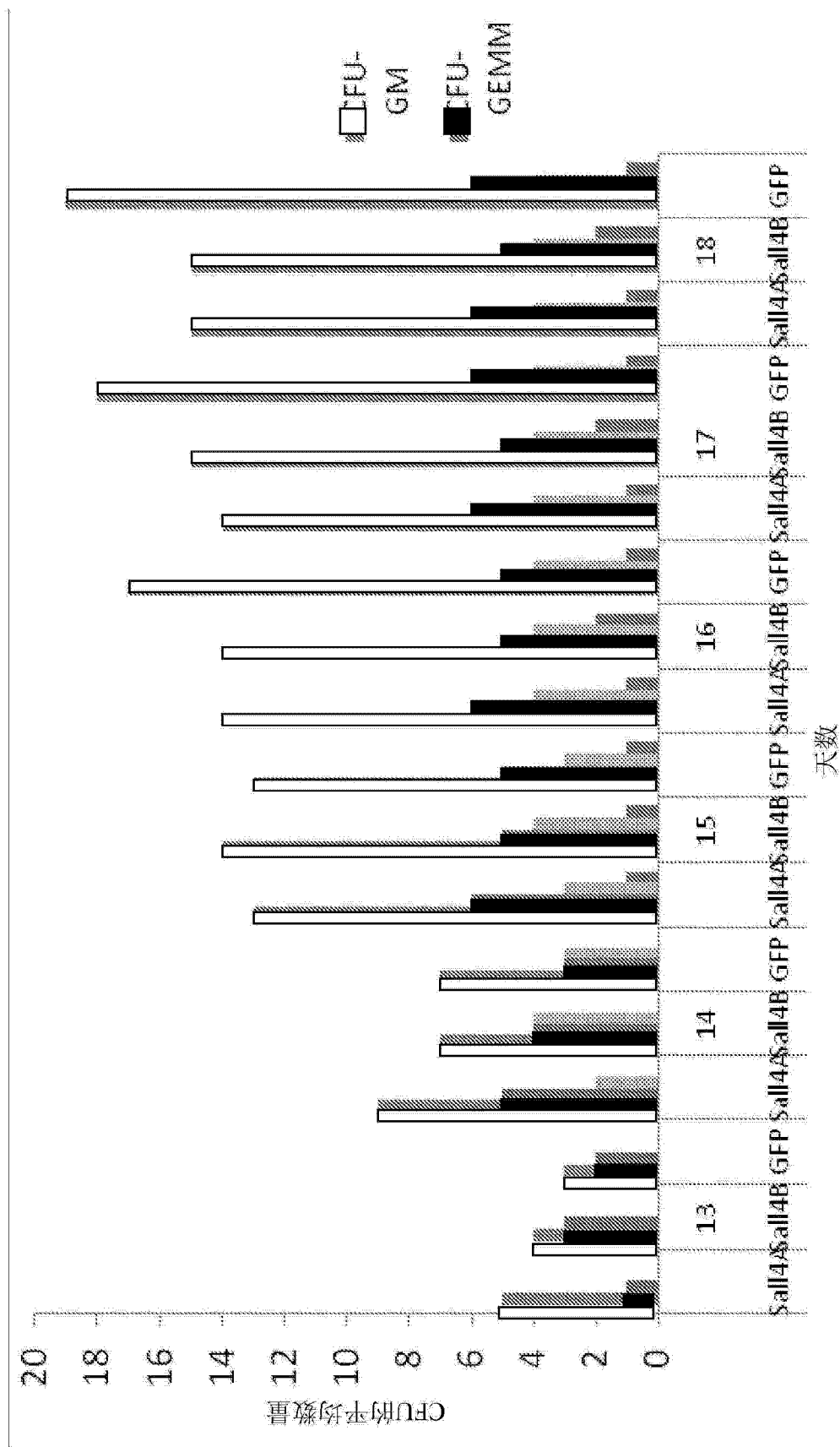


图 27B

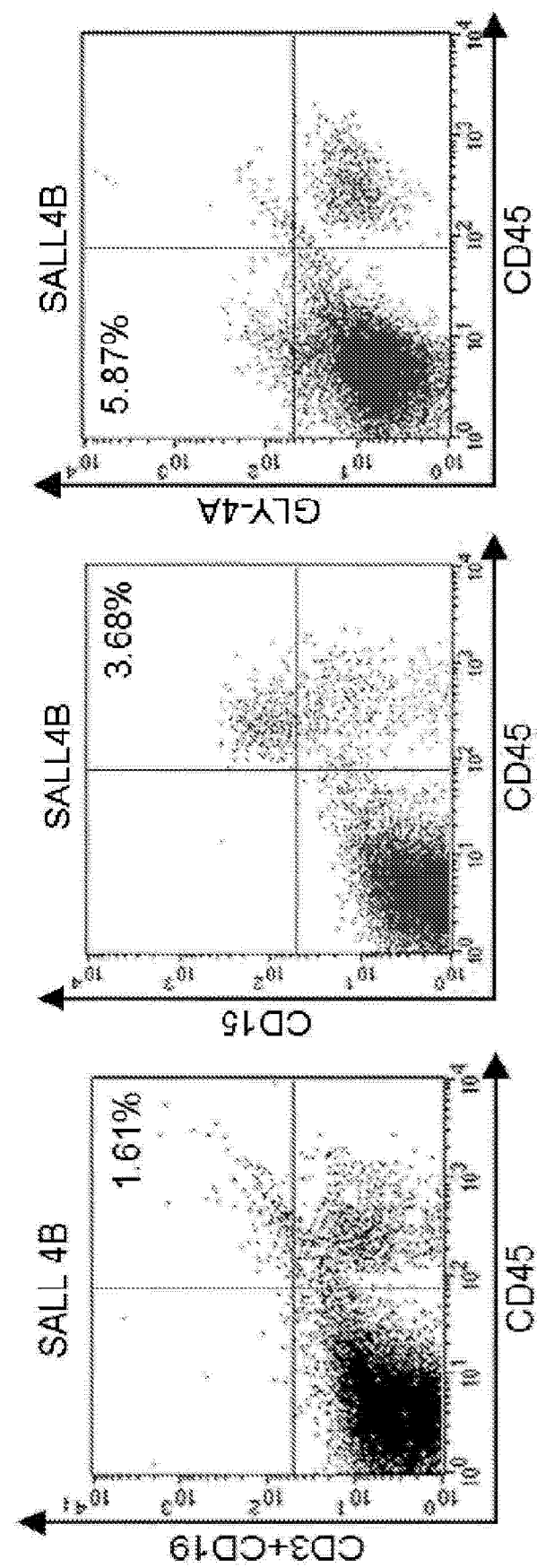


图 28

SALL4B蛋白质免疫印迹—SF9昆虫细胞

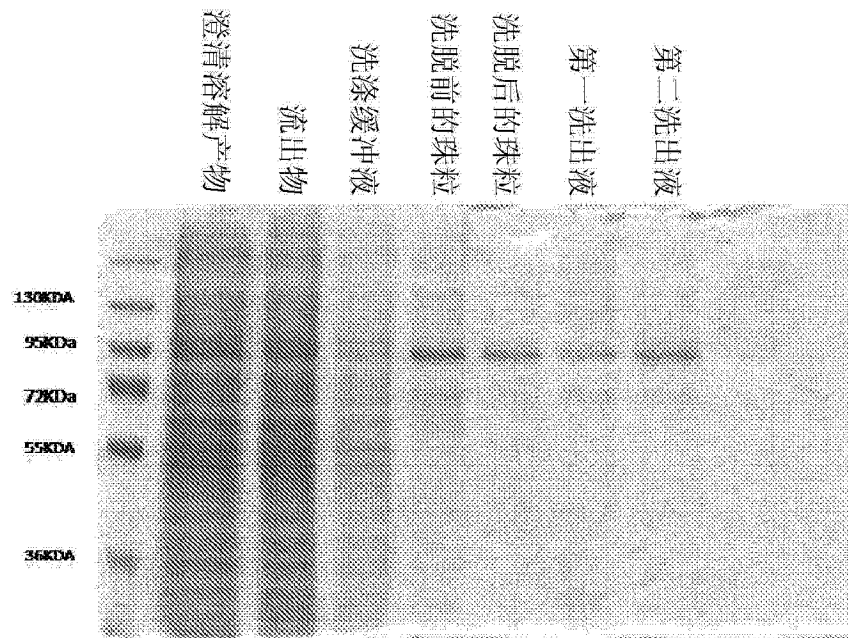


图 29

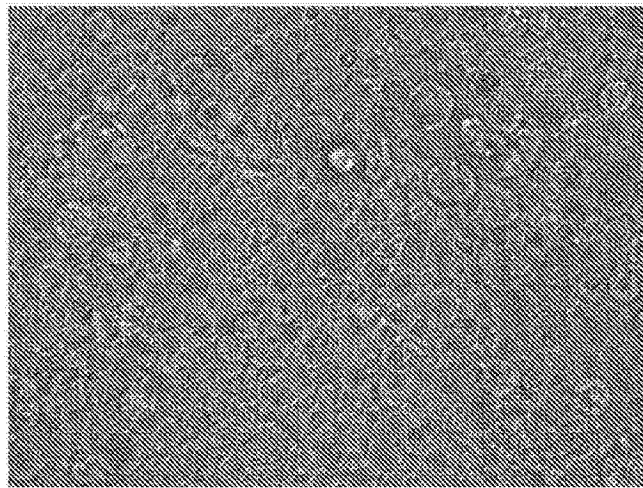


图 30

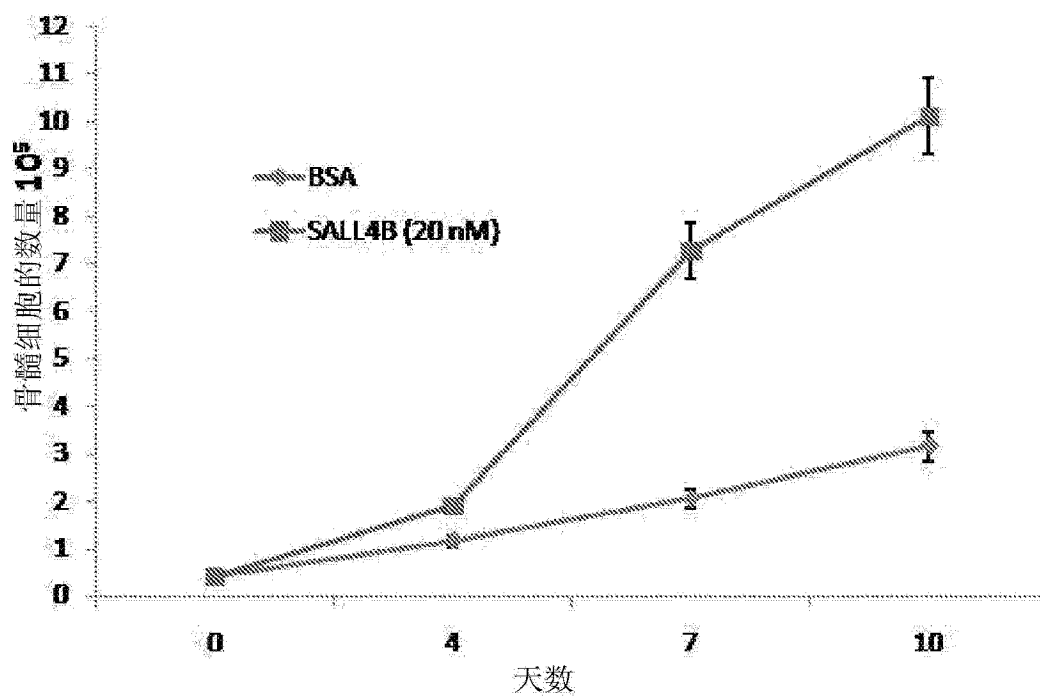


图 31

实验设计—使用SALL4B蛋白进行骨髓恢复或骨髓HSc/HPC的体内扩增

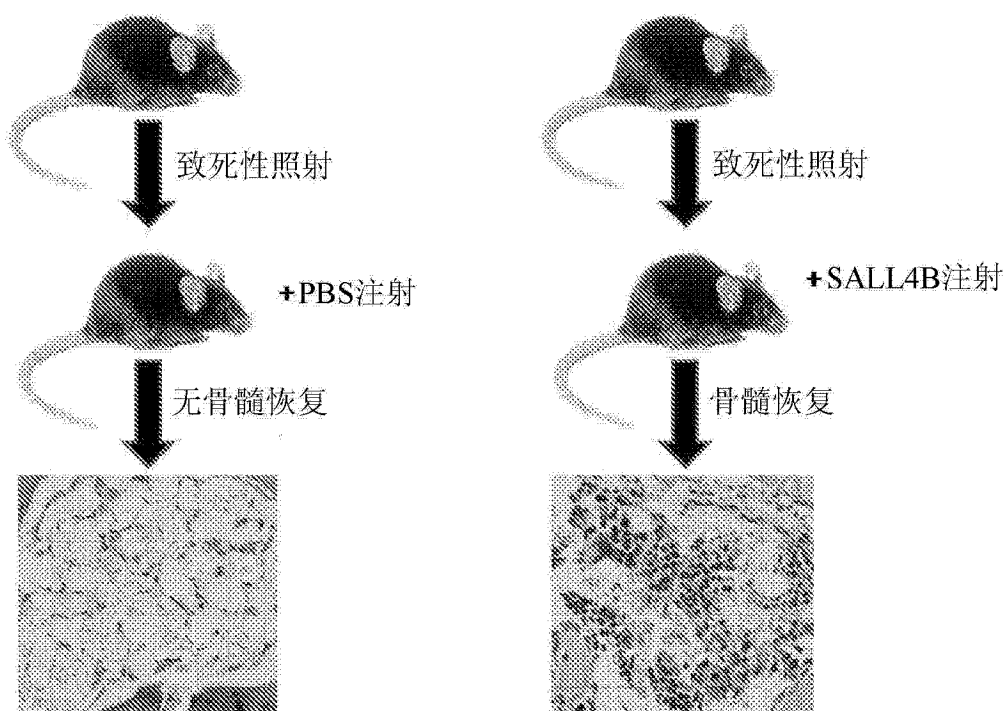


图 32

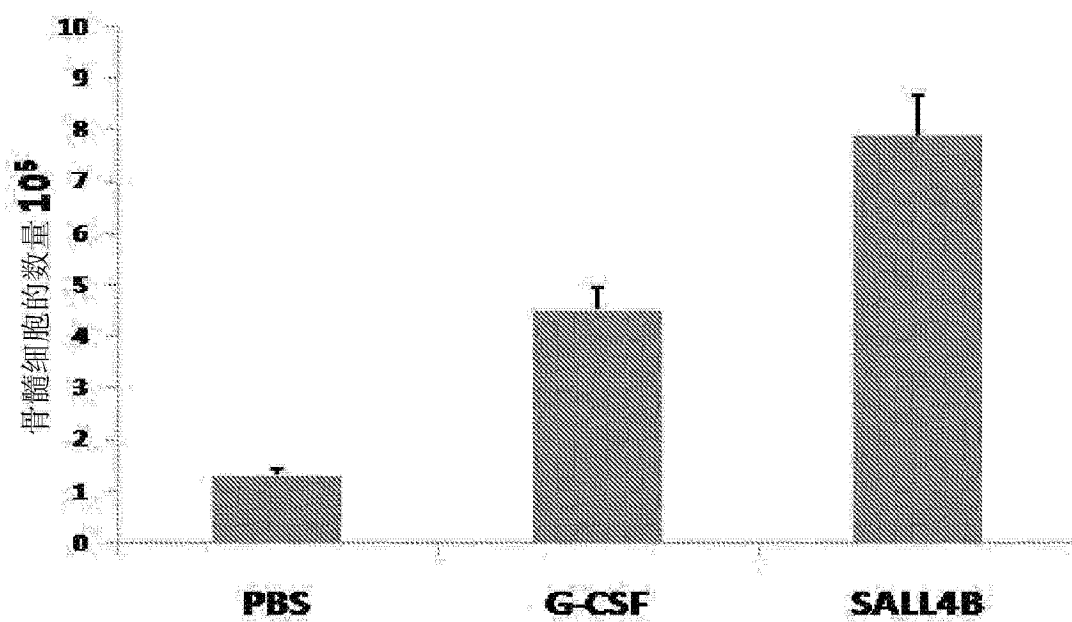


图 33

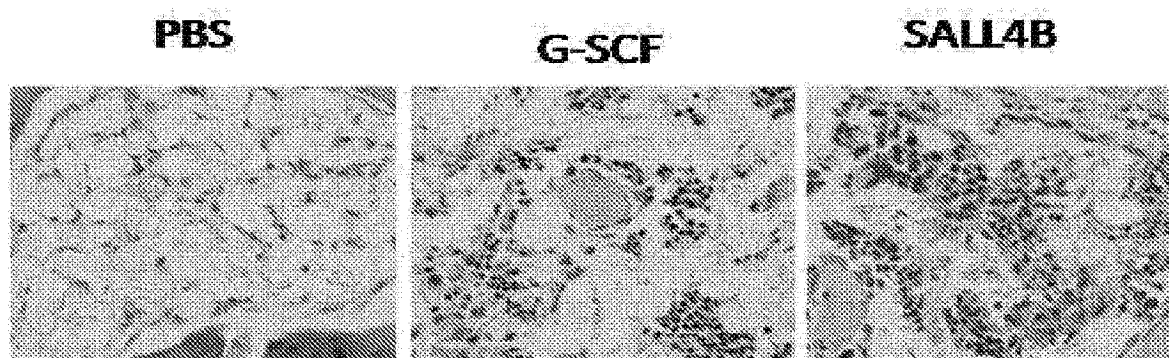


图 34

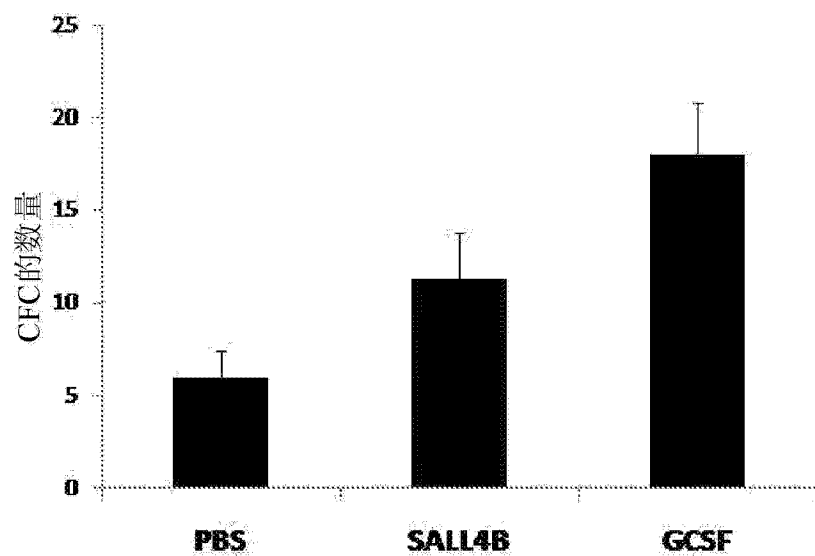


图 35

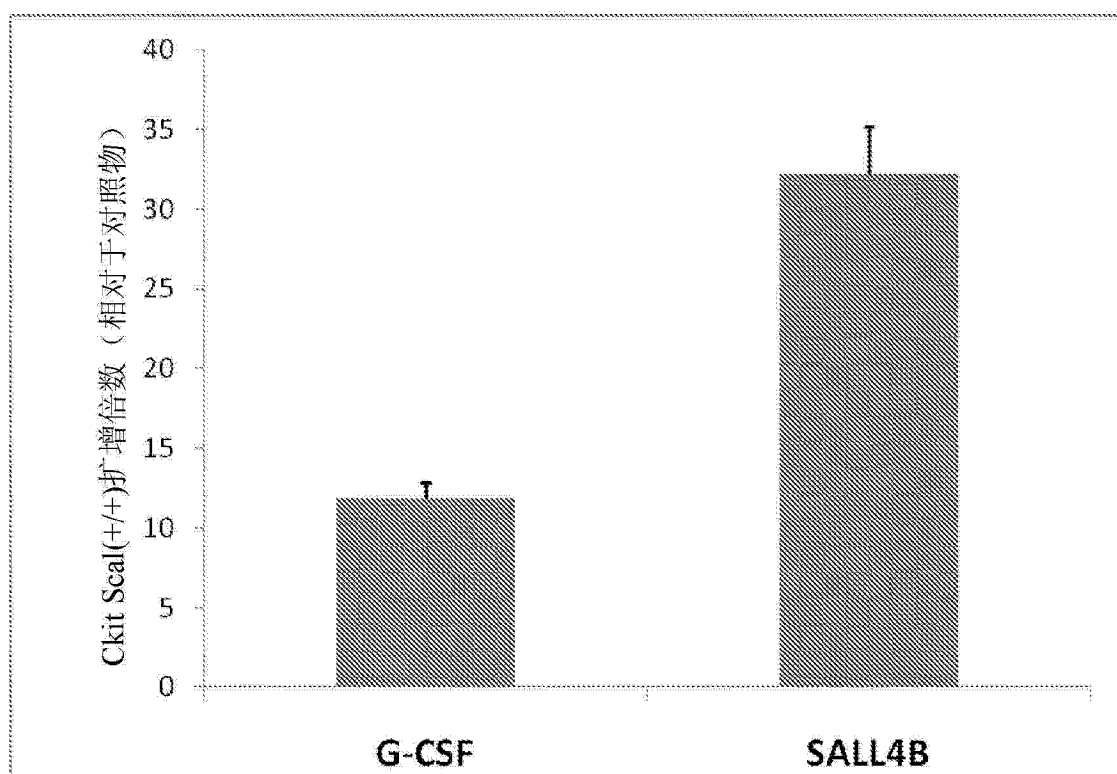


图 36

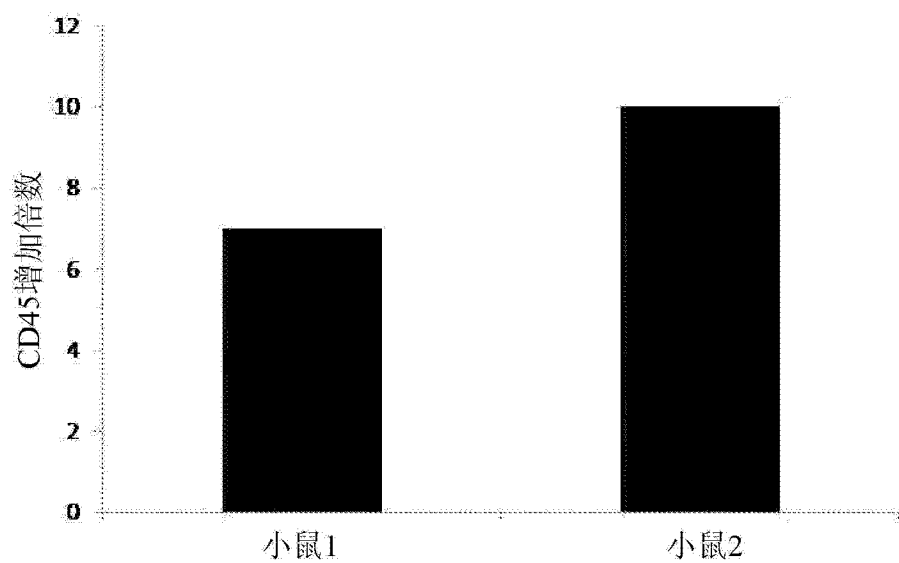


图 37

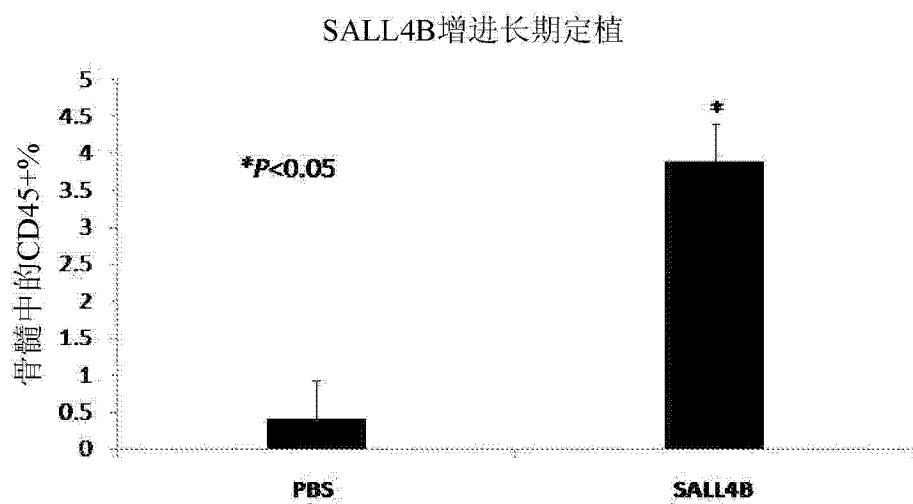


图 38

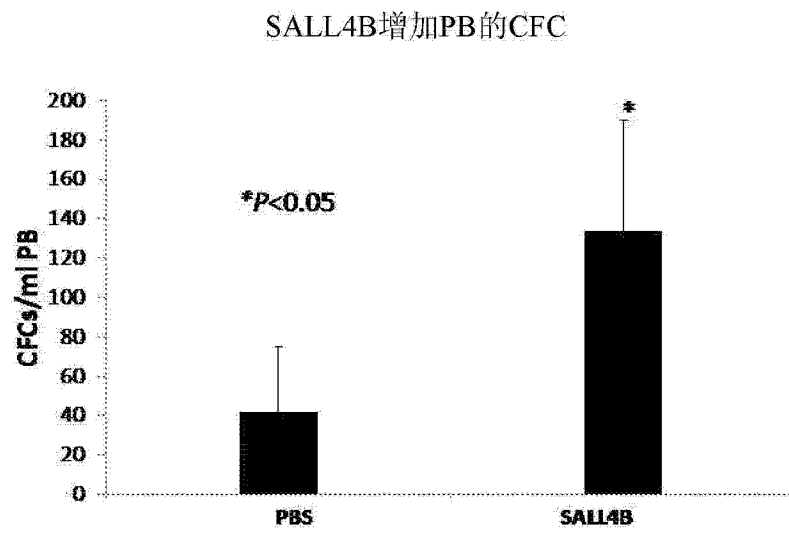


图 39